

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2004, 19:49:24 ; Search time 3021 Seconds  
(without alignments)  
3791.690 Million cell updates/sec

Title: US-09-513-365A-1  
Perfect score: 1481  
Sequence: 1 MLGQQQQQLYSALLTGER.....DNEKTMDSKSTKTKDRSR 280

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlp  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cn2\_1/USPTO\_spool\_p/US09513365/runat\_12012004\_163932\_28772/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513365 @CN 1 1 4958 @runat\_12012004\_163932\_28772 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*

29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrt: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1481	100.0	840	6	AR213554 Sequence
2	1481	100.0	1078	6	AR213555 Sequence
3	1481	100.0	1078	9	AB012853 Homo sapi
4	1481	100.0	1080	6	AX211560 Sequence
5	1481	100.0	1080	9	AF053537 Homo sapi
6	1481	100.0	1141	9	BC030128 Homo sapi
7	1478	99.8	843	9	HSAB6851 Homo sapi
8	1422.5	96.0	1262	10	BC050003 Mus muscu
9	1399.5	94.5	1001	10	AF078834 Mus muscu
10	1334	90.1	1154	6	AR220840 Sequence
11	1197	80.8	974	9	HSING252
12	1197	80.8	18317	9	AC107214 Homo sapi
13	1157	78.1	192753	2	AC107236 Homo sapi
14	1152	77.8	231311	2	AC132966 Rattus no
15	1152	77.8	245104	2	AC128348 Rattus no
16	1152	77.8	247796	2	AC094539 Rattus no
17	1152	77.8	255408	2	AC129658 Rattus no
18	1041	70.3	1094	5	AY014017 Xenopus l
19	1008	68.1	42029	2	AC114762 Homo sapi
20	962	65.0	666	4	AB098919 Bos tauru
21	880.5	59.5	840	10	AF149820 Mus muscu
22	874.5	59.0	1584	10	BC016573 Mus muscu
23	874.5	59.0	2817	10	AF177757 Mus muscu
24	853.5	57.6	840	9	AF078835 Homo sapi
25	853.5	57.6	845	9	HSAB10392 Homo sapi
26	853.5	57.6	873	6	AR171884 Sequence
27	853.5	57.6	911	9	AF044076 Homo sapi
28	853.5	57.6	1533	6	AX367042 Sequence
29	853.5	57.6	1533	9	AF149721 Homo sapi
30	853.5	57.6	2444	9	AB024401 Homo sapi
31	853.5	57.6	2897	9	AF181850 Homo sapi
32	750	50.6	1944	10	AF177756 Mus muscu
33	750	50.6	1959	10	AF177755 Mus muscu
34	749	50.6	184918	2	AC114608 Mus muscu
35	749	50.6	235079	2	AC124475 Mus muscu
36	748	50.5	201302	2	AC120896 Rattus no
37	748	50.5	222955	2	AC135534 Rattus no
38	733.5	49.5	888	9	AB037594 Homo sapi
39	730	49.3	1074	9	AF149722 Homo sapi
40	728	49.2	1143	6	AX367043 Sequence
41	728	49.2	1189	9	AF149723 Homo sapi
42	728	49.2	1902	6	AR079046 Sequence
43	728	49.2	1902	6	AR087457 Sequence
44	728	49.2	1902	6	AR110646 Sequence
45	728	49.2	1902	6	AR154563 Sequence

ALIGNMENTS

RESULT 1

```
AR213554      AR213554      840 bp      DNA      linear      PAT 25-SEP-2002
LOCUS          Sequence 6 from patent US 6403785.
ACCESSION      AR213554
VERSION        AR213554.1  GI:23310821
KEYWORDS       Unknown.
ORGANISM       Unknown.
REFERENCE      Unclassified.
AUTHORS        Nagata,M., Ozaki,K., Shimada,Y. and Horie,M.
TITLE          Isolated DNA molecule encoding human TSC403
JOURNAL        Patent: US 6403785-A 6 11-JUN-2002;
FEATURES       Location/Qualifiers
               source
               1..840
               /organism="unknown"
BASE COUNT     297 a 164 c 211 g 168 t
ORIGIN
1  MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
2  ATGTTAGGCGCAGCAGCAGCACTGTAAGTCTCGCGCGCGCTCTGACCGGGAGCGG 60
3  SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40
4  AGCGCGCTGCTCACTGCTACGTCGAGGAGTACCTTGAAGTGGAGTGGCTGCCAC 120
5  AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
6  GACATGCAGAGGAGCTGCTGCTGCGGAGAGTGGACAAATATCAAGAAACGTTA 180
7  LysGluLeuLeuAspValTyrGluLysTyrLysLysGluAspAsnGlnLysLys 80
8  AGGAAATTTGATGATGCTACGAAATATAGAAAGAGATGATTTAAACCAAGAGAA 240
9  ArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlnLeuGlyAspGluLys 100
10 CGTCTACAGCAGCTTCTCCAGAGAGCCTTAATTAATAGTCAAGAAATGGGAGATGAAAA 300
11 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
12 ATACAGATTTTACACAAATGCTCGAATTTGGTGGAAAAATCGGGCAAGACAAATGGAGTTA 360
13 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
14 CACTACAGTGTTCACAGATCTCTGCTGAAGTGAACGAGCTCAGATTAAGCAAGATG 420
15 AspSerSerGlnProGluArgSerArgProArgProArgGlnArgGlnThrSerGluSer 160
16 GATTCAGCAACCAAGAGATCTTCAAGAGAGACCCCGCAGCGAGCGAGTGAAGC 480
17 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProLysGlu 180
18 CTGATTTATGTCACATGGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAGAA 540
19 LysLysSerLysSerAlaLysLysLysArgSerLysAlaLysGlnGlnArgGluAla 200
20 AGAAATCCAGTCAGCAAGAGAAAGAAACGCTCCAGGCCACGCGAAGAGGAGCT 600
21 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
22 TCACCTGTTGAGTTTGCATATAGATCTTAATGAACCTACATATCTGCTTATGCAACCAAGTG 660

221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240
23 TCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATGAATGGTTTCACTTT 720
24 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260
25 TCATGTGTTTCACTTACCTATAAACCAAGGGGAAATGTAATGCCCAAGTGCAGGGGA 780
26 AspaGlnGlyLeuThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280
27 GATAATGAGAAACAATGCAAAAGTACTGAAAAAGACAAAAAGATAGATAGATCGAGG 840

RESULT 2
LOCUS          AR213555      1078 bp      DNA      linear      PAT 25-SEP-2002
DEFINITION     Sequence 7 from patent US 6403785.
ACCESSION      AR213555
VERSION        AR213555.1  GI:23310822
KEYWORDS       Unknown.
ORGANISM       Unknown.
REFERENCE      Unclassified.
AUTHORS        Nagata,M., Ozaki,K., Shimada,Y. and Horie,M.
TITLE          Isolated DNA molecule encoding human TSC403
JOURNAL        Patent: US 6403785-A 7 11-JUN-2002;
FEATURES       Location/Qualifiers
               source
               1..1078
               /organism="unknown"
BASE COUNT     356 a 206 c 280 g 236 t
ORIGIN
1  MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
2  ATGTTAGGCGCAGCAGCAGCACTGTAAGTCTCGCGCGCGCTCTGACCGGGAGCGG 151
3  SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40
4  AGCGCGCTGCTCACTGCTACGTCGAGGAGTACCTTGAAGTGGAGTGGCTGCCAC 211
5  AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
6  GACATGCAGAGGAGCTGCTGCTGCGGAGAGTGGACAAATATCAAGAAACGTTA 271
7  LysGluLeuLeuAspValTyrGluLysTyrLysLysGluAspAsnGlnLysLys 80
8  AGGAAATTTGATGATGCTACGAAATATAGAAAGAGATGATTTAAACCAAGAGAA 331
9  ArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlnLeuGlyAspGluLys 100
10 CGTCTACAGCAGCTTCTCCAGAGAGCCTTAATTAATAGTCAAGAAATGGGAGATGAAAA 391
11 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
12 ATACAGATTTTACACAAATGCTCGAATTTGGTGGAAAAATCGGGCAAGACAAATGGAGTTA 451
13 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
14 CACTACAGTGTTCACAGATCTCTGCTGAAGTGAACGAGCTCAGATTAAGCAAGATG 511
15 AspSerSerGlnProGluArgSerArgProArgProArgGlnArgGlnThrSerGluSer 160
16 GATTCAGCAACCAAGAGATCTTCAAGAGAGACCCCGCAGCGAGCGAGTGAAGC 571
```

QY 161 ArgAspLeuCysHisMetAlaSerGlyIleGluAspCysAspAspGlnProProLysGlu 180  
 DB 572 CGTGATTATGTGACATGCGCAATGGATTGAAGACTGTGATGATCAGCCACCTAAAGAA 631

QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnLysGluArgGluAla 200  
 DB 632 AAGAAATCCAGTTCAGCAAGAAAGAAAGAAACGCTCCAAGGCCAAGCAGGAAAGGAAGCT 691

QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220  
 DB 692 TCACCTGTGAGTTGCAATAGATCTCTAATGAACCTACATATCTGCTTATGCAACCAAGTG 751

QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240  
 DB 752 TCTTATGGGAGATGATAGATGTGACAAATGAACAGTGTCCAATGAATGGTTTTCACITT 811

QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260  
 DB 812 TCATGTGTTTCACTTACCTATTAACCAAGGGGAAATGGTATGCCCCAAGTGCAGGGGA 871

QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280  
 DB 872 GATAATGAGAAACAATGGCAAAAGTACTGAAAGACAAAAGAGATAGAGATCGAGG 931

## RESULT 3

LOCUS AB012853 1078 bp mRNA linear PRI 18-MAR-1999  
 DEFINITION Homo sapiens INGIL mRNA for INGILp, complete cds.

ACCESSION AB012853

VERSION AB012853.1 GI:4115554

KEYWORDS INGILp.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS Shimada,Y., Saito,A., Suzuki,M., Takahashi,E. and Horie,M.

TITLE Cloning of a novel gene (INGIL) homologous to ING1, a candidate

tumor suppressor

Cytogenet. Cell Genet. 83 (3-4), 232-235 (1998)

JOURNAL 99172097

MEDLINE 10072587

PUBMED 2 (bases 1 to 1078)

REFERENCE Shimada,Y., Saito,A. and Horie,M.

AUTHORS Direct Submission

TITLE Submitted (07-APR-1998) Yoshikazu Shimada, Otsuka Pharmaceutical

Co. Ltd., Otsuka GEN Research Institute, Kagasuno, Kawauchi-cho,

Tokushima, Tokushima 771-0192, Japan

(E-mail:shim@otsuka.genome.ad.jp, Tel:81-886-65-2888(ex.2476),

Fax:81-886-37-1035)

Sequence updated (17-Apr-1998).

## COMMENT

## FEATURES

## source

1..1078  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 /map="4q35.1"  
 /dev\_stage="fetal"  
 /tissue\_lib="fetal-brain"

## gene

1..1078

/gene="INGIL"

92..934

/gene="INGIL"

/codon\_start=1

/product="INGILp"

/protein\_id="BA36419.1"

/db\_xref="GI:4115555"

/translation="MWGQQQQLYSSAALLTGERSLTTCYVQDYLCVESLPHDMQR

NVSLRLDNKYETLEIDVDYKKEEDLNKKRLQQLRALINSLGDEKIQ

IVTWMLELVENRQMELHSCQFDPAESRASDKAKMSSQPERSPRRQRITSES

RDLCHMANGIEDDDPPKEKSKSAKKKRSKAKQEREASVPEAIDNPETCYLCN

QVSYGEMTIGCDNEQCIWFHFSCVSLTYKPKGKWYCPKCRGDNKTKMDKSTKTKKD  
 RRSR"

BASE COUNT 356 a 206 c 280 g 236 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.92e-113 Length: 1078  
 Score: 1481.00 Matches: 280  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-513-365A-1 (1-280) x AB012853 (1-1078)

QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluAcg 20  
 DB 92 ATGTTAGGGCAGCAGCAGCACTGCTCGCGCGCGCTCTGACCGGGAGCGG 151

QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValIcuSerLeuProHis 40  
 DB 152 AGCCGGCTGCTCACCTGCTACGTGCAGGACTACCTTGAGTGGTGGAGTGGCTGCCAC 211

QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60  
 DB 212 GACATGCAGAGGAACGTGCTGTGCTGCGAGAGCTGGACAACAAATATCAAGAAACGTTA 271

QY 61 LysGluIleAspAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLys 80  
 DB 272 AAGGAAATTTGATGATGCTACGAAATATTAAGAAAGAAAGATGATTTAAACCCAGAGAAA 331

QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlnGluGlyAspGluLys 100  
 DB 332 CGTCTACAGAGCTTCTCCAGAGACCACTAATTATAGTCAAGANTTGGAGATGAAAAA 391

QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120  
 DB 392 ATACAGATTGTTACAAATGCTGAAATGGTGGAAATCGGGCAAGACAAATGGAGTTA 451

QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140  
 DB 452 CACTCACAGTGTTCCTCAAGAGTCTGCTGAAAGTGAACGAGCTCAGATAAACCAAGATG 511

QY 141 AspSerSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSer 160  
 DB 512 GATTCACGCCAACGAAAGATCTTCAAGAGACCCCGCAGCAGCGGACCAAGTGAAGC 571

QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180  
 DB 572 CGTGATTTATGTACATGGCAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAA 631

QY 181 LysLysSerLysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200  
 DB 632 AAGAAATCCAGTTCAGCAAGAAAGAAAGAAACGCTCCAAGGCCAAGCAGGAAAGGAAGCT 691

QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220  
 DB 692 TCACCTGTGAGTTGCAATAGATCTCTAATGAACCTACATATCTGCTTATGCAACCAAGTG 751

QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240  
 DB 752 TCTTATGGGAGATGATAGATGTGACAAATGAACAGTGTCCAATGAATGGTTTTCACITT 811

QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260  
 DB 812 TCATGTGTTTCACTTACCTATAAACCAAGGGGAAATGGTATGCCCCAAGTGCAGGGGA 871

QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280  
 DB 872 GATAATGAGAAACAATGGCAAAAGTACTGAAAGACAAAAGAGATAGAGATCGAGG 931

## RESULT 4

AX211560

Percent Similarity: 100.00% Conservative: 0







```

Db      601 TCACCTGTTGAGTTTGCATAGATCTCTAATGAACCTACATACCTGCTTATGCAACCAAGTG 660
Qy      221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrPheHisPhe 240
Db      661 TCTTATGGGAGATGATAGATGTGACATGACATGACATGCTCCATTAATGTTTCACTTT 720
Qy      241 SerCysValSerLeuThrTyrIlysProIysGlyIleTyrTyrCysProIysCysArgGly 260
Db      721 TCATGTGTTTCACTTACCTATAAACCAAGGGGAAATGGTATTGCCAAAGTCAGGGGA 780
Qy      261 AspAsnGluThrThrMetIlePheSerThrGluIysThrIlysThrIlysAspArgSerArg 280
Db      781 GATAATGAGAAACAATGGCAAAAGTACTGAAAGACCAAAAAAGGATGATGAGATCGAGG 840

RESULT 8
BC050003
LOCUS   1262 bp mRNA linear ROD 01-APR-2003
DEFINITION Mus musculus, similar to inhibitor of growth family, member 1-like, clone IMAGE:6515125, mRNA, partial cds.
ACCESSION BC050003
VERSION   BC050003.1 GI:29436709
KEYWORDS
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 1262)
TITLE    Strausberg R.
JOURNAL  Direct Submission
Submitted (31-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT  Contact: MGC help desk
        Email: cgapsb@mail.nih.gov
        Tissue Procurement: Gilbert Smith, Ph.D.
        cDNA Library Preparation: Life Technologies, Inc.
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
        Web site: http://www.nisc.nih.gov/
        Contact: nisc_mgc@hghri.nih.gov
        Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 108 Row: 0 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 12963706.
Location/Qualifiers
    1..1262
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="FVB/N-3"
    /db_xref="taxon:10090"
    /clone="IMAGE:6515125"
    /tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."
    /clone_lib="NCI CGAP Mam2"
    /lab_host="DH10B"
    /notes="Vector: pCMV-SPORT6"
    425..>1262
    /codon_start=1
CDS

```

```

/product="Similar to inhibitor of growth family, member 1-like"
/protein_id="AAH50003.1"
/db_xref="GI:29436710"
/translation="MLGQQQQQQLYSSAALLTGERSLLTTCYVDLYECVESLPHDMQ
RNVSVLRDNRYQETLKEIDVYIKKEDDSNOKRLQQLQHLQRLALINSQELGDEKI
QIVTQMLELVENRARMELHSCFQDPAESERASDKSKMDSQPERSRPRRQRTSE
SRDLCHMTNGIDDCDDQPPKEXRSKAKKRSKAKQERASPVFAIDPNEPTYCLC
.NOVSGEMIGICDNEQCPIWFHFSVSLTYKPKGWYCPKRGDNKMTDKSTETKK
KKK"
BASE COUNT 320 a 310 c 422 g 210 t
ORIGIN
Alignment Scores:
Pred. No.: 1-61e-108 Length: 1262
Score: 1422.50 Matches: 269
Percent Similarity: 98.21% Conservatives: 5
Best Local Similarity: 96.42% Mismatches: 4
Query Match: 96.05% Indels: 1
DB: 10 Gaps: 1
US-09-513-365A-1 (1-280) x BC050003 (1-1262)
Qy      1 MetLeuGly---GlnGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGlu 19
Db      425 ATGTTAGGCGCAGCAGCAGCAGCAGCTGTACTCGTCGCCGCGCTCTGACCGGAGAG 484
Qy      20 ArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuPro 39
Db      485 CGAGCGCGCTCTCACCTGCTACCTGTCAGGAGTACTACCTGGAGTGTGTGGAGTGGTCCC 544
Qy      40 HisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnIysTyrGlnGluThr 59
Db      545 CACGACATGCAGAGGACGTCGTGCTGCGGGAGCTGGACACCAATATCCAGAAACG 604
Qy      60 LeuLysGluIleAspAspValTyrGluLysTyrIysLysGluAspAspLeuAsnGlnLys 79
Db      605 TTAAGGAAATTTGATGATGCTCTCAAAAATATAGAAAGAGATGATTTCCACACGAGAA 664
Qy      80 LysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlnLeuGlyAspGlu 99
Db      665 AAACCCCTACAGCAGCATCTCCAGAGAGCGTTTAATCAATAGCAAGAAATTTGGAGATGA 724
Qy      100 LysIleGlnIleValThrGlnMetLeuGluValGluAsnArgAlaArgGlnMetGlu 119
Db      725 AAAATTCAGATTGTCCACGAGATGCTCGAATTTGGTGGAGAACCGAGCAGACAAATGGAG 784
Qy      120 LeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLys 139
Db      785 CTGCATTACAGTGTTCACGATCTCTGCTGAAGTGAGCGAGCTCTCAGCAAGTCGAG 844
Qy      140 MetAspSerSerGlnProGluArgSerSerArgArgProArgProArgGlnArgThrSerGlu 159
Db      845 ATGGATTCCAGTCAACCGGAAAGATCTTCTAGAAGACCTTCGAAGACAGAGGACCATGAG 904
Qy      160 SerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLys 179
Db      905 ACCCGTGACTTAATGTCATGACAAACGAGGATTTGACACTGTGTGATGATCAACCCGAA 964
Qy      180 GluLysLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLysLysLysLys 199
Db      965 GAAAGAGATCCAGTCCGCCAAGAAAGAGAGCGCTCAAGGCCAAGCAGGAGGAG 1024
Qy      200 AlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGln 219
Db      1025 GCATCCCTGTCTGAGTTTGCCATCGATCCCAATGAGCCACCTACTGCTGTGTGTAAACAA 1084
Qy      220 ValSerTyrGlyGluMetIleGlyCysAspAsnGlnGlnCysProIleGluTyrPheHis 239
Db      1085 GTGTCTTACGGGGAGATGATAGGCTGTGACATGAAACAGTGTCCCATGTAATGGTTTCA 1144
Qy      240 PheSerCysValSerLeuThrTyrIlysProLysGlyLysTyrTyrCysProLysCysArg 259

```

```
Db 1145 TTTTCATGTGTTTCACTACCTATATAACCAACCGGGAATGGTATTGTCCTCAAAAGTGTAGG 1204
Qy 260 GlyAspAsnGluThrMetAspLysSerThrGluLysThrLysLysAspArgArg 278
Db 1205 GGAGACAATGAGAAAAACCATGGACAAAAGTACCGAAAAGACAAAAAAGAAAAA 1261

RESULT 9
AF078834
LOCUS AF078834 1001 bp mRNA linear ROD 09-SEP-2000
DEFINITION Mus musculus p33ING2 (Ing2) mRNA, complete cds.
ACCESSION AF078834
VERSION AF078834.1 GI:10039542
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1001)
Nagashima, M., Hagiwara, K., Hancock, A.R. and Harris, C.C.
Direct Submission
Submitted (16-JUL-1998) Laboratory of Human Carcinogenesis,
National Cancer Institute, National Institutes of Health, 37
Convent Dr., Bldg. 37, Rm. 2C01, Bethesda, MD 20892, USA
JOURNAL
FEATURES
Location/Qualifiers
1..1001
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..1001
/genes="Ing2"
8..853
/genes="Ing2"
/codon_start=1
/product="p33ING2"
/protein_id="AAG12173.1"
/db_xref="GI:10039543"
/translacion="MLGQQQQOOLYSAAALLTGRSRLTTCVQDYLCVSLPHDQ
RVNVLRLDNKYQETLKEIDVYKVEDDSNOKRLQHLQRLALINSOELGDEKI
QNTQMLGLVNRSMOELHSOCFQDPAESERASDKSMDSSQPRSRPRRQRTSE
SRDLCHMTNGIDDDQPPKREKRSKAKKREASVSFAIDNPFTYCLC
NQVSGEMIGCDNEQCPFIWFHSCVSLTYKHKRWYCPKCRGDNKMTDKSTKTK
ERROR"
BASE COUNT 324 a 205 c 264 g 208 t
ORIGIN
Alignment Scores:
Pred. No.: 9.49e-107 Length: 1001
Score: 1399.50 Matches: 266
Percent Similarity: 97.15% Conservative: 7
Best Local Similarity: 94.66% Mismatches: 7
Query Match: 94.50% Indels: 1
DB: 10 Gaps: 1

US-09-513-365A-1 (1-280) x AF078834 (1-1001)
Qy 1 MetLeuGly---GlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuThrGlyGlu 19
Db 8 ATGTTAGGCGAGCAGCAGCAGCAGCAGTACTCTCGCGCGCGCTCTCTGACCGGAGAG 67
Qy 20 ArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuPro 39
Db 68 CGGAGCGCGCTGCTACCTGCTACGTACGTGCGAGGACTACTGGAGTGTGTGGAGTGGCTGCC 127
Qy 40 HisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThr 59
Db 128 CACGACATGCAGAGGAACGTGTCGGTCTCGCGAGCTGCACAAACAAATACCAAGAAACG 187
Qy 60 LeuLysGluLeuAspAspValTyrGluLysTyrLysLysGluAspAsnGlnLys 79
Db 188 TTAAGGAAATTCATGATGCTATGAAAAATATGAGAAAAGAGATGATTCACCAACCAAGAA 247
Qy 80 LysArgLeuGlnGlnLeuLeuGlnArgAlaLeuAsnSerGlnGlnLeuGlyAspGlu 99
```

```
Db 248 AAACGCCTACAGCAGCATCTCCAGAGAGCGTTAATCAATAGCAAGAAATGGGAGATGAA 307
Qy 100 LysIleGlnIleValThrGlnMetLeuGluValGluAsnArgAlaArgGlnMetGlu 119
Db 308 AAAATTCAAGATGTACCACAGATCTCGGATTGGTGGAGAACCGATCGAGACAAATGGAG 367
Qy 120 LeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLys 139
Db 368 CTGCATTACAGTGTTCACAGATCTCTGTAAGTGGAGCGCTCAGACAAAGTCGAAG 427
Qy 140 MetAspSerSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGlu 159
Db 428 ATGGATTCCAGTCAACCGGAAAAGATCTTCTAGAACCTCGAAGACAGAGGACCAATGAG 487
Qy 160 SerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProLys 179
Db 488 AGCCGTGATCTTATGTACATGACAAACGGGATTGACGACTGTGATGATCAACCCGAAA 547
Qy 180 GluLysLysSerLysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGlu 199
Db 548 GAAAGAGATCCAAGTCCGCCAAGAGAGAGCGCTCCAAGGCCAAGCAGGAGGAG 607
Qy 200 AlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGln 219
Db 608 GCATCCTCTGTGAGTTCGATCCATCCCAATGAGCCACCTACTGCTTGTGTAACCAA 667
Qy 220 ValSerTyrGluMetIleGlyCysAspAsnGlnGlnCysProIleGluThrPheHis 239
Db 668 GTGTCTACGGGAGATGATAGGCTGTGACATGAACAGTGTCCCATTAATGATGTTTAC 727
Qy 240 PheSerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArg 259
Db 728 TTTTCATGTGTTTCACTCACTATATAACACAGGGGAAATGGTATTGCCCAAGTGTAGG 787
Qy 260 GlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSer 279
Db 788 GGAGACAATGAGAAAACCAATGCAAAAAGTACCGAAAAGACAAAAAAGGAGAGAGCG 847
Qy 280 Arg 280
Db 848 AGG 850

RESULT 10
AR220840
LOCUS AR220840 1154 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 81 from patent US 6426186.
ACCESSION AR220840
VERSION AR220840.1 GI:23327717
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1154)
AUTHORS Jones, K.A., Volkmut, W. and Walker, M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 81 30-JUL-2002;
FEATURES location/Qualifiers
1..1154
/organism="unknown"
BASE COUNT 367 a 216 c 285 g 265 t 21 others
ORIGIN
Alignment Scores:
Pred. No.: 2.95e-101 Length: 1154
Score: 1334.00 Matches: 270
Percent Similarity: 81.87% Conservative: 1
Best Local Similarity: 81.57% Mismatches: 9
Query Match: 90.07% Indels: 53
DB: 6 Gaps: 1

US-09-513-365A-1 (1-280) x AR220840 (1-1154)
Qy 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuThrGlyGluArg 20
```





```

repeat_region 3376. .4311 /rpt_family="MER1_type"
repeat_region 4312. .4642 /rpt_family="ERV1"
repeat_region 4643. .4762 /rpt_family="Alu"
repeat_region 4763. .5003 /rpt_family="ERV1"
repeat_region 5009. .5298 /rpt_family="MER1_type"
repeat_region 5299. .5605 /rpt_family="Alu"
repeat_region 5606. .6889 /rpt_family="MER1_type"
repeat_region 6690. .6920 /rpt_family="Alu"
repeat_region 6921. .7263 /rpt_family="MER1_type"
repeat_region 7405. .7452 /rpt_family="L2"
repeat_region 7793. .7864 /rpt_family="MER1_type"
repeat_region 8141. .8933 /rpt_family="L2"
repeat_region 8235. .8583 /rpt_family="L1"
repeat_region 8584. .8853 /rpt_family="Alu"
repeat_region 8854. .8933 /rpt_family="L1"
repeat_region 8953. .9114 /rpt_family="Alu"
repeat_region 9120. .9417 /rpt_family="Alu"
repeat_region 9422. .9836 /rpt_family="Alu"
repeat_region 9946. .10061 /rpt_family="Alu"
repeat_region 10696. .10958 /rpt_family="Alu"
repeat_region 10964. .11110 /rpt_family="MER1_type"
repeat_region 11580. .11886 /rpt_family="Alu"
repeat_region 12156. .12476 /rpt_family="Alu"
repeat_region 12488. .13030 /rpt_family="ERV1"
repeat_region 13077. .13192 /rpt_family="("TTCC)n"
repeat_region 13225. .13538 /rpt_family="Alu"
repeat_region 13634. .13749 /rpt_family="MIR"
repeat_region 14111. .14206 /rpt_family="L1"
repeat_region 14565. .14864 /rpt_family="Alu"
repeat_region 14982. .15277 /rpt_family="Alu"
repeat_region 15357. .15652 /rpt_family="Alu"
repeat_region 15709. .16001 /rpt_family="Alu"
repeat_region 16730. .16751 /rpt_family="Alu"
repeat_region 18157. .18271 /rpt_family="Alu"
repeat_region 19348. .19529 /rpt_family="Alu"
repeat_region 19874. .20012 /rpt_family="L1"
repeat_region 20068. .20103 /rpt_family="L1"

/rpt_family="AT_rich"
20263. .20564 /rpt_family="Alu"
20769. .20995 /rpt_family="Alu"
21095. .21383 /rpt_family="Alu"
21605. .21740 /rpt_family="Alu"
21746. .22060 /rpt_family="Alu"
22072. .22381 /rpt_family="Alu"
22432. .22526 /rpt_family="MIR"
23055. .23181 /rpt_family="Alu"
23182. .23212 /rpt_family="("TAAA)n"

Alignment Scores:
Pred. No.: 4,27e-87 Length: 183317
Score: 1197.00 Matches: 231
Percent Similarity: 93.98% Conservative: 3
Best Local Similarity: 92.77% Mismatches: 9
Query Match: 80.82% Indels: 6
DB: Gaps: 2

US-09-513-365A-1 (1-280) x AC107214 (1-183317)

QY 32 LeuGluCysValGluSerLeuProHisAspMetGlnArgAsnValSerValLeuArgGlu 51
Db 45149 TTGGAATGTTGCTGCTGCTAACACATGAT-----ACGTTCTCATTTTCTTTTC 45199

QY 52 LeuAspAsnLysTyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLys 71
Db 45200 CTT-----TTTTTGAACCGTTAAAGAAATTTGATGTTCTACGAAAAATATAAG 45250

QY 72 LysGluAspAspLeuAsnGlnLysLysArgLeuGlnLeuLeuGlnArgAlaLeuIle 91
Db 45251 AAAGAAGATGATTTAAACCAGAGAAACGCTTACAGACGCTTCTCCAGAGAGCACATAAT 45310

QY 92 AsnSerGlnGluLeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuVal 111
Db 45311 ATAGTCAGAAATTTGGAGATGAAAAAATACAGATTGTTTACACAAATGCTCGAATTTGGTG 45370

QY 112 GluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSer 131
Db 45371 GAAATCGGGCAAGACAAATGGAGTTACACTCACAGTGTTCACAGATCCTGCTGAAAGT 45430

QY 132 GluArgAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerSerArgArg 151
Db 45431 GAACGAGCCTCAGATAAAGCAAGATGGATTCCAGCCAAACCAAGAAAGATCTTCAAGAAGA 45490

QY 152 ProArgArgGlnArgThrSerGluSerArgAspLeuCysHisMetAlaAsnGlyIleGlu 171
Db 45491 CCCCCGAGCAGCGGACCGACCTGAAGCCGTGATTTATGTCAATGTCATGTCGCAATGGGANTGAA 45550

QY 172 AspCysAspAspGlnProProLysGluLysLysSerLysSerLysAlaLysLysLysArg 191
Db 45551 GACTGTGATGATCAGCCACCTTAAAGAAAAGAAATCCAACTCAGCAAGAAAAGAAACGC 45610

QY 192 SerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGlu 211
Db 45611 TCCAAGGCCAAGCAGGAAAGGGAAGCTTCACCTGTTGAGTTTGCATAGATATCTTAATGAA 45670

QY 212 ProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGlu 231
Db 45671 CCTACATCTGCTTATGCAACCAAGTGCTTATGGGGAGATGATAGGATGTGCAATGAA 45730

QY 232 GlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysGly 251
Db 45731 CAGTGTCCAATTTGAATGGTTTTCATTTTCATGTTTCTTACTTACCTATAACCAAGGGG 45790

```



QY 252 LysTTPtyrCysProLysCysArgGlyAspAenGluLysThrMetAspLysSerThrGlu 271  
 |||||||  
 Db 45791 AATGTTATTCCTCAAGTGCAGGGAGATATGAGAAACATGGACAAAGTACTGAA 45850  
 |||||||

QY 272 LysThrLysLysAspArgArgSerArg 280  
 |||||||

Db 45851 AAGACAAAAGAGTAGAAGATCGAGG 45877  
 |||||||

RESULT 13  
 AC107236/c  
 LOCUS  
 DEFINITION Mus musculus clone RP23-296C23, WORKING DRAFT SEQUENCE, 7 unordered  
 pieces.  
 AC107236  
 VERSION AC107236.4 GI:28876115  
 HTG; HTGS PHASE1; HTGS DRAFT.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Birren, B., Nusbaum, C. and Lander, E.  
 1 (bases 1 to 192753)  
 Mus musculus, clone RP23-296C23  
 Unpublished  
 2 (bases 1 to 192753)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgaiter, B.,  
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
 Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A.,  
 Cooke, P., DeAtellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
 Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,  
 Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,  
 MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
 Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission

TITLE  
 JOURNAL  
 Submitted (17-JAN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 192753)

REFERENCE  
 AUTHORS  
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
 Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,  
 Boguslavsky, L., Boukhgaiter, B., Canarata, J., Chang, J., Choepel, Y.,  
 Collamore, A., Cooke, A., Cooke, P., Corum, B., DeAtellano, K.,  
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,  
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,  
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B.,  
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
 Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,  
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
 Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,  
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
 Rachupka, A., Ramaeamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,  
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission

## JOURNAL

## COMMENT

Submitted (07-MAR-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 7, 2003 this sequence version replaced gi:28195970.  
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L19778

Center clone name: 296.C.23

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 191671 bases at least Q40

Consensus quality: 191880 bases at least Q30

Consensus quality: 192003 bases at least Q20

Insert size: 188000; agarose-fp

Insert size: 192153; sum-of-ctnigs

Quality coverage: 9.1 in Q20 bases; agarose-fp

Quality coverage: 8.9 in Q20 bases; sum-of-ctnigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 2597: contig of 2597 bp in length  
 \* 2598 2697: gap of 100 bp  
 \* 2698 5130: contig of 2433 bp in length  
 \* 5131 5230: gap of 100 bp  
 \* 5231 91253: contig of 86023 bp in length  
 \* 91254 91353: gap of 100 bp  
 \* 91354 102343: contig of 10990 bp in length  
 \* 102344 102443: gap of 100 bp  
 \* 102444 135243: contig of 32800 bp in length  
 \* 135244 135343: gap of 100 bp  
 \* 135344 190271: contig of 54928 bp in length  
 \* 190272 190371: gap of 100 bp  
 \* 190372 192753: contig of 2382 bp in length.

## FEATURES

## Source

1. 192753  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP23-296C23"  
 /clone\_lib="RPC1-23 Female Mouse BAC"

## misc\_feature

1. 2597  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left

## misc\_feature

2698..5130

/note="assembly\_fragment"

## misc\_feature

5231..91253

/note="assembly\_fragment"

## misc\_feature

91354..102343

/note="assembly\_fragment"

## misc\_feature

102444..135243

/note="assembly\_fragment"

## misc\_feature

135344..190271

/note="assembly\_fragment"

## misc\_feature

190372..192753

/note="assembly\_fragment"

clone\_end:T7  
 vector\_side:right

## TITLE

BASE COUNT 54639 a 42472 c 42123 g 52919 t 600 others

## ORIGIN

```

Alignment Scores:      9.21e-84      Length:      192753
Pred. No.:            1157.00
Score:                98.22%
Percent Similarity:    95.56%
Best Local Similarity: 78.12%
Query Match:          2
DB:                   0

US-09-513-365A-1 (1-280) x AC107236 (1-192753)

QY 56 TyrglnGluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspAsp 75
   ::
Db 98076 TTTTGAACACGTTAAAGGAATTTGATGTTCTATGAAAAATATAAGAAAGAGATGAT 98017

QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuLeuSerGlnGlu 95
   ::
Db 98016 TCCAAACGAGAAAAACGCCCTACAGCAGCATCTCCAGAGAGCGTTAATCAATAGCCCAAGAA 97957

QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115
   ::
Db 97956 TTGGGAGATGAAAAATTCAGATTGTCACCCAGATGCTCGAATTTGTGGAGAACCCGAGCG 97897

QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSer 135
   ::
Db 97896 AGACAAATGAGCTGCATTCACAGTGTTCACAGATCTCTCTAGAAAGCTCCAGAGCCTCA 97837

QY 136 AspiysAlaLysMetAspSerSerGlnProGluArgSerSerArgProArgArgGln 155
   ::
Db 97836 GACAAGTCGAAGATGATTCAGTCAACCGGAAGATCTTCTAGAAAGCTCCAGAGACAG 97777

QY 156 ArgThrSerGluSerArgAspLeuGluMetAlaAsnGlyIleGluAspCysAspAsp 175
   ::
Db 97776 AGACCACTGAGAGCGTGTGATTCATGTCATGACAAACGGGATTCAGCACTGTGATGAT 97717

QY 176 GlnProLysGluLysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLys 195
   ::
Db 97716 CAACCCGAGAAAGAGATCCAGTCCGCCCAAGAGAGAGCGCTCCAGGCCCAAG 97657

QY 196 GlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCys 215
   ::
Db 97656 CAGGAGAGGAGGAGCATCCCTGTCGATTTGCGCATCCATCCCAATGAGCCCACTACTGC 97597

QY 216 LeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIle 235
   ::
Db 97596 TTGTGTAAACCAAGTGTCTACGGGGAGATGATAGGCTGTGACATGACAGTGTCCCAT 97537

QY 236 GluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCys 255
   ::
Db 97536 GATGTGTTTCATTTTCATGTTGTTTCACTCACTATAACCCAGGGGAATGTATGTC 97477

QY 256 ProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLys 275
   ::
Db 97476 CCAAGGTGTAGGGAGACATGAGAAACCATGACAAAGATGCCGAAAGACAAAAAG 97417

QY 276 AspArgArgSerArg 280
   ::
Db 97416 GAGAGAAGACGCGAGG 97402

RESULT 14
AC132966/c
LOCUS AC132966 231311 bp DNA linear HTG 19-NOV-2002
DEFINITION Rattus norvegicus clone CH230-210B24, WORKING DRAFT SEQUENCE, 2
unnumbered pieces.
ACCESSION AC132966
VERSION AC132966.4 GI:25073360
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

REFERENCE  
AUTHORS

1 (bases 1 to 231311)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Amin, A., Anguiano, D., Allen, C., Allen, H., Alsbrooks, S., Ayodeji, M., Baca, E., Baden, H., Anyalebechi, V., Aoyagi, A., Barber, M., Barnstead, M., Benahmed, F., Baldwin, D., Bandaranaike, D., Barker, M., Blyth, P., Brown, M., Biswal, K., Blair, J., Blankenburg, K., Burch, P., Burrell, K., Calderon, E., Bryant, N., Buhay, C., Carter, K., Cavazos, I., Ceasar, H., Center, A., Cardenas, V., Carter, K., Chen, R., Chen, Y., Chen, Z., Chu, J., Chacko, J., Chavez, D., Chen, G., Cox, C., Coyle, M., Cree, A., D'Souza, L., Cleveland, C., Cockrell, R., Davy, C., De Anda, C., Dederich, D., Davila, M. L., Davis, C., Deramo, C., Ding, Y., Dinh, H., Divya, K., Delgado, O., Denson, S., Dunn, A., Durbin, K., Duval, B., Faves, K., Draper, H., Dugan-Rocha, S., Evans, C., Falls, T., Fan, G., Egan, A., Escotto, M., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fernandez, S., Finley, M., Flagg, N., Garcia, A., Garner, T., Garza, M., Fraser, C. M., Gabisi, A., Ganta, R., Grady, K., Gill, R., Guerra, W., Guevara, W., Gebregregis, E., Geer, K., Hladun, S. L., Hodgson, A., Hoques, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Hawlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Johnson, B., Johnson, R., Jolivet, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kovar, C., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kratt, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, H., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasiermak, S., Paul, H., Perez, A., Perez, L., Frankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodargren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlaczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE  
JOURNAL

2 (bases 1 to 231311)

Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (05-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 231311)

Rat Genome Sequencing Consortium.  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information

Center project name: KBZ1

Center clone name: CH230-210B24

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 215506 bases at least Q40

Consensus quality: 218204 bases at least Q30

Consensus quality: 220107 bases at least Q20

Estimated insert size: 224360; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 225634: contig of 225634 bp in length

\* 225635 225734: gap of unknown length

\* 225735 231311: contig of 5577 bp in length.

Location/Qualifiers

1. 231311

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-210B24"

3446.102459

/notes="clone\_boundary"

clone\_end:T7

site:

end sequence: BH343917"

complement(218103..218757)

/notes="clone\_boundary"

clone\_end:Sp6

site:

end sequence: BH343918"

20260..222134

/notes="wgs\_end\_extension"

clone\_end:Sp6"

222185..225634

/notes="wgs\_end\_extension"

clone\_end:Sp6"

clone\_end:Sp6"

BASE COUNT 60763 a 49448 c 50231 g 60699 t 10170 others

# ORIGIN

## Alignment Scores:

Pred. No.:	3,02e-83	Length:	231311
Score:	1152.00	Matches:	214
Percent Similarity:	98.22%	Conservative:	7
Best Local Similarity:	95.11%	Mismatches:	4
Query Match:	77.79%	Indels:	0
DB:	2	Gaps:	0

US-09-513-365A-1 (1-280) x AC132966 (1-231311)

Qy 56 TrrGlnGluThrLeuLysGluLeuAspValTrrGluLysTrrLysLysGluAsp 75

Db 53296 TTTTGAACCTTTAAAGGAATTGATGCTCTATGAAAAATATAAGAAAGATGAT 53237

Qy 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlu 95

Db	53236	TCAAACCCAGAAAAACGCTTACAGCAGCATCTCCAGAGCAGCATTATCAATAGCCAAAGAA	53177
Qy	96	LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla	115
Db	53176	TTGGGAGATGAAAGATTTCAGATTTCACACAGATGCTGGGAATTGGTGGAAAAACCGAGCC	53117
Qy	116	ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerArgGluArgAlaSer	135
Db	53116	AGACAAATGGAGCTGCATTCACAGTGTTCACAGGATCTCTGCTGAAGTGAGCGGCGCTCA	53057
Qy	136	AspLysAlaLysMetAspSerSerGlnProGluArgSerArgProArgProArgGln	155
Db	53056	GACAAATCCAGATGGATTCCAGTCAACCCGAAAGATCTTCCAGAAAGACCTCGAAGACAG	52997
Qy	156	ArgThrSerGluSerArgAspLeuGluCysHisMetAlaAsnGlyIleGluAspCysAspAsp	175
Db	52996	CGGACACAGTGAGCGCGTATTATGTGTACATGACAAACGGGATTGATGACGCGATGAT	52937
Qy	176	GlnProLysGluLysLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLys	195
Db	52936	CAGCCACCGAAGAAAGAAATCCAAATCCGCCAAGAAAGAAAGCGCTCCAAAGGCCAAG	52877
Qy	196	GlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGluProThrTrrCys	215
Db	52876	CAGGAGAGGAGCGCTCACCTGTTGAGTTTGCATAGATCCCAACGAGCCACCTACTGTC	52817
Qy	216	LeuCysAsnGlnValSerTrrGlyGluMetIleGlyCysAspAsnGluGlnCysProfile	235
Db	52816	TTATGTAAACCAAGTGTCTTACGGGAGATGATAGGCTGTGCAATGAACAGTGTCCCAT	52757
Qy	236	GluTrrPheHisPheSerCysValSerLeuThrTrrLysProLysGlyLysTrrTrrCys	255
Db	52756	GAGTGGTTTCACTTTCTGCTTTCCTCTACCTATAACCAAGGGAATGGTATTGCT	52697
Qy	256	ProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLys	275
Db	52696	CCAAAGTCAGGGGAGACAGTGAGAAACCATGGACAAAGTACTGMAAAGACAAAGAAAG	52637
Qy	276	AspArgArgSerArg 280	
Db	52636	GAGAGAAGCGGAGG 52622	

## RESULT 15

AC128348/c

LOCUS

DEFINITION

AC128348

AC128348

AC128348.3

GI:25007900

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Rattus norvegicus

ORGANISM

REFERENCE

1 (bases 1 to 245104)

AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anylebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaruaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, K., Chen, Y., Chen, Z.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunarathne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Krat, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Loulsegged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Ngunidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nuyuen, N., Norris, S., Nwaokelemeh, O., Okwuonu, G., Olarnpunaagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sned, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Stealm, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

# TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Unpublished  
2 (bases 1 to 245104)  
Worley, K.C.

Direct Submission  
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245104)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23265010.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: G2CX

Center clone name: CH230-129P3

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 202567 bases at least Q40

Consensus quality: 204753 bases at least Q30

Consensus quality: 205914 bases at least Q20

Estimated insert size: 206745; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 241682: contig of 241682 bp in length  
\* 241683 241782: gap of unknown length  
\* 241783 242944: contig of 1162 bp in length  
\* 242945 243044: gap of unknown length  
\* 243045 245104: contig of 2060 bp in length.

## FEATURES

### source

1..245104  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-129P3"

### misc\_feature

1..1042  
/note="wgs end extension  
clone\_end:Sp6"

### misc\_feature

2108..3020  
/note="clone boundary  
clone\_end:Sp6"

### misc\_feature

240971..241683  
/note="clone\_boundary  
clone\_end:17  
site:  
end sequence: BH310917"

BASE COUNT 57525 a 45075 c 46272 g 57482 t 38750 others  
ORIGIN

### Alignment Scores:

Pred. No.: 3.25e-83 Length: 245104  
Score: 1152.00 Matches: 214  
Percent Similarity: 98.22% Conservative: 7  
Best Local Similarity: 95.11% Mismatches: 4  
Query Match: 77.73% Indels: 0  
DB: 2 Gaps: 0

US-09-513-365A-1 (1-280) x AC128348 (1-245104)

Qy 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrlsGluAsp 75  
Db 111137 TTTTAGAACTTTAAAGAAATTGATGTCATGAAATAATAAGAAAGATGAT 111078  
Qy 76 LeuAenGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlu 95  
Db 111077 TCACACAGAAAAACGGCTACAGAGCATCTCCAGAGGCATTATCAATAGCCAGAA 111018  
Qy 96 LeuGlyAspGluLysLleGlnlleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
Db 111017 TTGGAGATGAAAGATTTCAGATTTCACACAGATGCTGGAATTGGTGGAAACCCAGCC 110958  
Qy 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSer 135  
Db 110957 AGACAAATGGAGCTGCATTTCAGTGTTCAGAGTCTCGTGAAGTGAAGCGGCTCA 110898  
Qy 136 AspLysAlaLysMetAspSerSerGlnProGluArgSerSerArgProArgArgGln 155  
Db 110897 GACAAATCCAGATGATTCAGTCCAGCCGGAAGATCTTCAGAGACCTCAAGACAG 110838  
Qy 156 ArgThrSerGluSerArgAspLeuGlnCysHisMetAlaAsnGlylleGluAspCysAsp 175  
Db 110837 CGGACCACTGAGAGCGGTGATTATGTCACATGACAAACGGGATGATGACTGCGATGAT 110778

```
QY      176  GlnProProLysGluLysLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLys 195
Db      110777  CAGCCACCGAAGAAAGAAATCCAAATCGCCNAGAAAGAAAGCGCTCCAGGCCAAG 110718

QY      196  GlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCys 215
Db      110717  CAGGAGAGGGAAGCGTCACCTGTTGAGTTTGCCATAGATCCCAACGAGGCCACCTACTGC 110658

QY      216  LeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIle 235
Db      110657  TTATGTAAACCAAGTGCTTTACGGGGAGATGATAGGCTGTGACAAATGAACAGTGTCCTCCATT 110598

QY      236  GluTtpPheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTrpTyrCys 255
Db      110597  GAGTGGTTTCACTTTTCTCGCGTTTTCACCTACCTATAAACCAAGGGGAAATGGTATTGC 110538

QY      256  ProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLys 275
Db      110537  CCAAAAGTCAGGGGAGACAGTGAAGAAACCATGGACAAAGTACTGAAAGACAGAAAG 110478

QY      276  AspArgArgSerArg 280
Db      110477  GAGAGAAGAGCGAGG 110463
```

Search completed: January 12, 2004, 22:12:19  
Job time : 3230 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2004, 17:23:57 ; Search time 268 Seconds  
(without alignments)  
2820.308 Million cell updates/sec

Title: US-09-513-365A-1  
Perfect score: 1481  
Sequence: 1 MLCQQQQQLYSSAALLTGER.....DNEKTDWKSTKDKRRSR 280

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09513365/runat\_12012004\_163931\_28766/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_19Jun03 -QFW=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFW=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513365@cgn 1.1 490 @runat\_12012004\_163931\_28766 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1481	100.0	840	20	AAZ08595	Human INGLL encodi
2	1481	100.0	1078	20	AAZ08596	Human INGLL encodi
3	1481	100.0	1080	22	AAZ12783	Tumour suppressor
4	1345	90.8	1153	21	AAZ47473	Human tumour suppr
5	1334	90.1	1154	24	ABS70424	Human bone remodel
6	874.5	59.0	2817	21	AAZ53790	Murine P37ING1 cod
7	853.5	57.6	911	21	AAZ53792	Human P37ING1 codi
8	853.5	57.6	1533	22	AAH28478	Nucleotide sequenc
9	853.5	57.6	2897	24	ABK86977	Human inhibitor of
10	745	50.3	1835	21	AAZ53789	Murine ingl common
11	728	49.2	1143	22	AAH28479	Nucleotide sequenc
12	728	49.2	1902	18	AAZ69651	Tumour suppressor
13	728	49.2	1902	19	AAV62285	Partial INGL parti
14	728	49.2	1905	24	AAZ46126	Human tumour suppr
15	728	49.2	2061	18	AAZ69652	Tumour suppressor
16	728	49.2	2061	19	AAV62292	Human INGL full-le
17	728	49.2	2061	20	AAZ28688	Nucleotide sequenc
18	728	49.2	2886	24	ABK86976	Human inhibitor of
19	728	49.2	8487	22	AAZ37089	Human musculoskele
20	728	49.2	8487	25	ABX60077	cDNA encoding nove
21	728	49.2	163350	24	AAZ46127	Human tumour suppr
22	604	40.8	742	22	AAH28480	Nucleotide sequenc
23	602	40.6	857	22	AAH28481	Nucleotide sequenc
24	572	38.6	346	21	AAZ28459	Human secreted pro
25	480	32.4	693	24	ABQ50058	Oligonucleotide fo
26	480	32.4	693	24	ABQ50059	Oligonucleotide fo
27	466	31.5	693	24	ABQ50060	Oligonucleotide fo
28	466	31.5	693	24	ABQ50061	Oligonucleotide fo
29	455	30.7	1705	22	AAK51611	Human polynucleoti
30	449	30.3	1082	25	ABQ77429	Human CGDD cDNA 72
31	447	30.2	1772	23	ABL12805	Drosophila melanog
32	447	30.2	1781	21	AAZ53979	p33 tumour suppress
33	446.5	30.1	1465	21	AAZ15008	cDNA encoding a hu
34	446.5	30.1	1611	22	ABA09175	Human tumour suppr
35	446.5	30.1	1611	22	AAK52595	Human polynucleoti
36	436.5	29.5	987	23	ABL15055	Drosophila melanog
37	432.5	29.2	1365	24	ABL54076	Human tumour suppr
38	427.5	28.9	1108	22	AAZ69039	Cell cycle protein
39	424.5	28.7	958	22	AAZ69040	Cell cycle protein
40	424.5	28.7	1300	24	ABZ11521	Human polynucleoti
41	406	27.4	1465	22	AAK52101	Human polynucleoti
42	400	27.0	699	21	AAZ27422	IkappaB kinase (IK
43	396.5	26.8	678	21	AAZ27423	IkappaB kinase (IK
44	392.5	26.5	807	22	AAZ69040	Cell cycle protein
45	392.5	26.5	807	22	AAZ69040	Cell cycle protein

ALIGNMENTS

RESULT 1  
AAZ08595  
ID AAZ08595 standard; cDNA; 840 BP.  
XX  
AC AAZ08595;  
XX  
AC  
DT 18-OCT-1999 (first entry)  
XX  
XX Human INGLL encoding cDNA.  
DE  
XX  
XX Human; TSC403; INGLL; diagnosis; lung cancer; cell cycle; regulation;  
KW cell proliferation; cell aging; apoptosis; tumour suppressor; ss.  
XX  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT CDS 1..840

```
FT      /*tag= a
FT      /product= "INGIL"
FT      /note= "no stop codon given"
XX
PN      WO9940190-A1.
XX
PD      12-AUG-1999.
XX
PF      02-FEB-1999; 99WO-JP00419.
XX
PR      28-APR-1998; 98JP-0134679.
PR      03-FEB-1998; 98JP-0038133.
PR      05-MAR-1998; 98JP-0073234.
XX
PA      (SAKA ) OTSUKA PHARM CO LTD.
XX
PI      Horie M, Nagata M, Ozaki K, Shimada Y;
XX
DR      WPI; 1999-494294/41.
DR      P-PSDB; AAY29606.
XX
PT      Human lung-specific gene TSC430 overexpressed in cancer tissue, used
PT      for treatment of, e.g. colon tumour
XX
PS      Claim 14; Page 91; 99pp; Japanese.
XX
CC      The present sequence represents the human tumour suppressor gene INGIL.
CC      The present invention also describes the human gene TSC403 expressed
CC      specifically in normal lung tissue. TSC403 is useful in the
CC      diagnosis, investigation and treatment of cancers in which it is
CC      overexpressed, including cancer of the lung, breast, fallopian tube,
CC      oesophagus, colon, thyroid, parotid gland, bladder, ovary or pancreas.
CC      INGIL is useful in the investigation of cell proliferation, aging and
CC      apoptosis and the pathology of cancer, the diagnosis and treatment of
CC      cancer such as cancer of the colon, stomach, oesophagus or fallopian
CC      tube, and the screening of candidate drugs for the treatment of such
CC      cancers.
XX
SQ      Sequence 840 BP; 297 A; 164 C; 211 G; 168 T; 0 other;

Alignment Scores:
Pred. No.:      6.16e-133      Length:      840
Score:          1481.00      Matches:      280
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              20      Gaps:        0

US-09-513-365A-1 (1-280) x AAZ08595 (1-840)
QY      1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
Db      1 ATGTTAGGGCAGCAGCAGCAGCAACTGTACTCGTCGCCGCCGCTCCTCGACCGGGAGCG 60
QY      21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40
Db      61 AGCCGGCTGCTACCTGCTACGTCAGGACCTACCTTGAGTGGCTGGAGTCGTCGCCAC 120
QY      41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
Db      121 GACATGACAGGAACTGCTGCTGCTGCGAGAGCTGGACACAAATATCAAGAAAGCTTA 180
QY      61 LysGluIleAspAspValTyrGluLysTyrLysGluAspAspLeuAsnGlnLysLys 80
Db      181 AAGGAAATTCATGATGCTACGAAAAATATAGAAAGAAAGATGATTTAAACCAAGAGAAA 240
QY      81 ArgLeuGlnGlnLeuGlnArgAlaLeuIleAsnSerGlnLeuGluLeuGlyAspGluLys 100
Db      241 CGTCTACAGCAGCTTCTTCAGAGAGCAGCACTAATTAATAGTCAAGAAATGGGAGATGAAA 300
QY      101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db      301 ATACAGATTGTTACACAAATGCTCGAATGGTGGAAAAATCGGCAAGACAAATGGAGTTA 360
```

```
QY      121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
Db      361 CACTTCACAGTGTTTCAAGATCTCTGCTGAAAGTGAAGAGCTCAGTAAGCAAGATG 420
QY      141 AspSerSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSer 160
Db      421 GATTCCAGCCCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGCGAGCCAGTGAAGC 480
QY      161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspGlnProProLysGlu 180
Db      481 CGTGATTTATGTACATGGCAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAA 540
QY      181 LysLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLysLysLysLys 200
Db      541 AAGAAATCCAGTCAGCAAGAAAGAAACCTCAAGGCCCAAGCAGCAAGAAAGGAGCT 600
QY      201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db      601 TCACCTGTTGAGTTTGAATAGATCTTAATGAACCTACATACCTGCTTATGCAACCAAGTG 660
QY      221 SerTyrGlyGluMetIleGlyCysAspAsnGlnGlnCysProIleGluThrPheHisPhe 240
Db      661 TCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATGGAATGGTTTCACTTT 720
QY      241 SerCysValSerLeuThrTyrLysProLysGlyLysLysLysLysLysLysLysLys 260
Db      721 TCATGTGTTTCACTTACCTATTAACCAAGGGGAAATGGTATTTGCCAANGTCAGGGGA 780
QY      261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg 280
Db      781 GATATGAGAAACAAATGACAAAGTACTGAAAGACAAAAGAGATAGAGATCGAGC 840

RESULT 2
AAZ08596
ID      AAZ08596 standard; cDNA; 1078 BP.
XX
AC      AAZ08596;
XX
DT      18-OCT-1999 (first entry)
XX
DE      Human INGIL encoding cDNA with 5' and 3' untranslated regions.
XX
KW      Human; TSC403; INGIL; diagnosis; lung cancer; cell cycle; regulation;
KW      cell proliferation; cell aging; apoptosis; tumour suppressor; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      92..934
FT      /*tag= a
FT      /product= "INGIL"
XX
PN      WO9940190-A1.
XX
PD      12-AUG-1999.
XX
PF      02-FEB-1999; 99WO-JP00419.
XX
PR      28-APR-1998; 98JP-0134679.
PR      03-FEB-1998; 98JP-0038133.
PR      05-MAR-1998; 98JP-0073234.
XX
PA      (SAKA ) OTSUKA PHARM CO LTD.
XX
PI      Horie M, Nagata M, Ozaki K, Shimada Y;
XX
DR      WPI; 1999-494294/41.
DR      P-PSDB; AAY29606.
XX
PT      Human lung-specific gene TSC430 overexpressed in cancer tissue, used
PT      for treatment of, e.g. colon tumour
XX
```





```

QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40
DB 128 AGCCGGCTGCTACCTGCTACGTCAGGACTACCTTGAGTGGTGGAGTCGCTGCCAC 187

QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
DB 188 GACATCGACAGGAACGTGCTGTGCTGCGAGAGCTGGACACAAATATCAAGAAACGTTA 247

QY 61 LysGluLeuAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLys 80
DB 248 AAGGAAATTCATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCCAGAGAAA 307

QY 81 ArgLeuGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlnLysLysGluAspGluLys 100
DB 308 CGTCTACAGAGCTTCTCAGAGAGCAGCAATTAATAGTCAAGAAATGGAGATGAAAA 367

QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
DB 368 ATACAGATTGTTACACAAATGCTCGAATTTGGTGAATAATCGGCAAGACAAATGGAGTTA 427

QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
DB 428 CACTCACAGTGTTCAGAGATCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATG 487

QY 141 AspSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSer 160
DB 488 GATTCAGGCAACCAAGAAAGATCTTCAGAAAGACCCCGCAGGCGGACCAAGTGAAGC 547

QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180
DB 548 CGTGATTATGTCACATGCGCAATGGATTGAAGACTGTGATGATCAGCACCTAAAGAA 607

QY 181 LysLysSerLysSerLysLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
DB 608 AAGAAATCCAAGTCAGCAAGAAAGAAACGCTCCAAAGGCCCAAGCAGGAAAGGAAGCT 667

QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
DB 668 TCACCTGTGTGAGTTTGCATAGATCTCTTAATGAACCTACATACCTGTTATGCAACCAAGTG 727

QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGlnGlnCysProIleGluTrpPheHisPhe 240
DB 728 TCTTATGGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTCGAATGGTTTCACTTT 787

QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260
DB 788 TCATGTGTTCATTACCTATTAACCAAGGGGAAATGGTATTGCCCAAGTCAGGGGA 847

QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280
DB 848 GATAATGAGAAACAATAGTGAACAAAGTACTGAAAGACAAAAGGATAGATCGAGG 907

```

## RESULT 4

AAZ47473

ID AAZ47473 standard; cDNA; 1153 BP.

XX AC AAZ47473;

XX DT 14-MAR-2000 (first entry)

XX DE Human tumour suppressor (TUSUP) nucleotide sequence.

XX KW Tumour suppressor protein; TUSUP; human; cancer; treat; prevent;

XX OS reproductively tract; gastrointestinal tract; immune system; ss.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT 277..999

XX FT /\*tag= a

XX FT /product= TUSUP

XX FT /note= "tumour suppressor protein"

```

XX WO9961612-A1.
XX 02-DEC-1999.
XX 20-MAY-1999; 99WO-US11136.
XX 28-MAY-1998; 98US-0086359.
XX (INCY-) INCYTE PHARM INC.
XX Tang YT, Corley NC, Patterson C;
XX WPI: 2000-062711/05.
XX P-PSDB; AAY52199.
XX New human tumour suppressor protein for treating cancer, particularly of
XX reproductively and gastrointestinal tracts or immune system -
XX Claim 7; Fig 1; 71pp; English.
XX
XX This sequence encodes the human tumour suppressor (TUSUP) protein. The
XX TUSUP sequence was identified from a human lung cDNA library. The protein
XX is 240 amino acids in length and has one potential CAMP and cGMP
XX dependent protein kinase phosphorylation site, nine potential protein
XX kinase C phosphorylation sites, five potential casein kinase
XX phosphorylation sites, and one potential tyrosine kinase phosphorylation
XX site. TUSUP has structural and chemical similarity to p33. TUSUP can be
XX used to treat or prevent cancer, particularly of the reproductive or
XX gastrointestinal tracts or of the immune system. Agonists of the TUSUP
XX protein can be used to raise TUSUP specific antibodies and to screen for
XX specific binding agents and potential therapeutic agents. The TUSUP
XX nucleotide sequence can be used in gene therapy and as a source of
XX diagnostic probes and primers.
XX
XX Sequence 1153 BP; 367 A; 215 C; 285 G; 265 T; 21 other;

```

```

Alignment Scores:
Pred. No.: 1,08e-119 Length: 1153
Score: 1345.00 Matches: 270
Percent Similarity: 82.12% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 9
Query Match: 90.82% Indels: 52
DB: 21 Gaps: 1

```

US-09-513-365A-1 (1-280) x AAZ47473 (1-1153)

```

QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaLeuLeuLeuThrGlyGluArg 20
DB 9 ATGTTAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 68

QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40
DB 69 AC-CGGCTGCTCACCTGCTACGTGCGAGACTACCTTGAGTGGTGGAGTCGCTGCCAC 127

QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln----- 57
DB 128 GACATCGACAGAGAACGTGCTGTGCTGCGAGA-CTGGACACAAATATCAAGGTCTGGGC 186

QY 57 ----- 57
DB 187 TGTGGGGGGCGTGTGTTGGCCCCCAGCGGAGTCCGAATCGGGTTTGCAGCATGTTT 246

QY 57 ----- 57
DB 247 CGGTGATGTTTCCAACTCTTTCCAGTCAATGGATCAGGACGGCATCAGCAGCTCGGA 306

QY 58 -----GluThrLeuLysGluIleAspValTyrGluLysTyr 70
DB 307 CGGTGGGATCTCTGGCTCCGCAACGTTAAGGAAATTTGATGTCTTACGAAAAATAT 366

QY 71 LysLysGluAspAspLeuAsnGlnLysLysArgLeuGlnGlnLeuLeuArgAlaLeu 90

```



Db 727 ACCTCAAGGCCAAGCAGGAAGGAGCTTACCTGTTGAGTTTGCATATGATCCTAA 786  
 QY 210 nGluProThrTyrCysLeuGlnValSerTyrGlyGluMetIleGlyCysAspAs 230  
 Db 787 TGAACCTACATACCTGCTTATGCAACCAAGTCTTATGGGAGATGATAGGATGACAA 846  
 QY 230 nGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyrIlysProLy 250  
 Db 847 TGAACAGTGTCCAAATGATGGTTTCATTTTCATGTGTTTCACTTACCTATTAACCA 906  
 QY 250 sGlyLysThrTyrCysProLysCysArgGlyAspAsnGluLysThrMetAspLysSerTh 270  
 Db 907 GGGCAATGCTATTGCCCCAAGTGCAGGGGAGATATGAAAAACAATGCAAAAAGTAC 966  
 QY 270 rGlyLysThrLysAspArgSerArg 280  
 Db 967 TGAAGAAGACAAAAGGATAGATCGAGG 997

## RESULT 6

AAA53790  
 ID AAA53790 standard; cDNA; 2817 BP.

XX AAA53790;

DT 19-DEC-2000 (first entry)

XX Murine P37ING1 coding sequence.

DE p53; tumour; cancer; detection; antibody; hybridisation; exon 1b;  
 KW INGI1; INGI1; p37ING1; p37ING1; oncogene; gene therapy; diagnosis;  
 KW proliferation disorder; transformation; transfection; transformed cell; mouse; ds.  
 XX Mus musculus.

XX Key Location/Qualifiers  
 FH CDS 847..1686  
 FT /\*tag= a  
 FT /product= P37ING1 polypeptide

XX WO200046370-A1.  
 XX 10-AUG-2000.  
 XX 04-FEB-2000; 2000WO-US02959.  
 XX 04-FEB-1999; 99US-0118941.  
 XX (UNII ) UNIV ILLINOIS FOUND.

XX Gudkov A, Zeremski M, Gurova KV, Grigorian IA;  
 XX WPI; 2000-491278/43.  
 DR P-PSDB; AAY97242.

XX Detecting nucleic acid encoding exon 1b of INGI1, useful for diagnosing  
 PT and treating cancer, comprises contacting sample with isolated nucleic  
 PT acid comprising sequence of exon 1b and detecting hybridized products  
 XX Disclosure; Fig 11; 134pp; English.

XX Mutations in or loss of the p53 gene occur in more than 50% of  
 CC human tumours and tumour cell lines, but functional inactivation of  
 CC the p53 pathway occurs in a much larger proportion of tumours. In  
 CC many cases the mechanism of functional inactivation of the p53 gene  
 CC remains unknown but p53 has been found to act in cooperation with  
 CC INGI1. Functional cooperation between INGI1 and p53 suggested that  
 CC INGI1 encoded a tumour suppressor protein that functioned within the  
 CC p53 pathway. This data suggested a possible role for INGI1 in head  
 CC and neck cancers and chromosomal location of the INGI1 placed it  
 CC within a region that is frequently rearranged in head and neck  
 CC cancers. Large scale analysis of tumours involving INGI1 has not  
 CC revealed mutations in INGI1 nor significant variations in its  
 CC expression suggesting that INGI1 was not a useful gene to study in

CC cancer etiology. However, alternative initiation exons of the INGI1  
 CC gene, each having their own promoter have been discovered.  
 CC Expression of one promoter (1a) produces a protein identical to  
 CC INGI1. Expression of a second promoter (1b) produces a protein having  
 CC an identical C-terminal fragment to INGI1 but an additional 104  
 CC N-terminal amino acids. The newly discovered protein has been  
 CC designated p37ING1 (wild type: p33ING1). p37ING1 has the  
 CC characteristics of an oncogene. When overexpressed in cells (even  
 CC those expressing wild type p53) p37ING1 is able to cause  
 CC proliferation or transformation of those cells. Thus detecting a  
 CC nucleic acid encoding exon 1b of INGI1 by hybridisation with an  
 CC isolated nucleic acid having the sequence of exon 1b of INGI1  
 CC or its antisense sequence can identify individuals expressing the  
 CC oncogenic form of INGI1. Novel peptide sequences taken from the 104  
 CC N-terminal peptide of p37ING1 can also be used to raise antibodies  
 CC that can also be used in detection methods for the p37ING1 variant.  
 CC The polypeptides may be useful in gene therapy for treatment of cell  
 CC proliferation disorders, especially cancers and for diagnosing and  
 CC studying cancers.

XX SQ Sequence 2817 BP; 670 A; 711 C; 846 G; 590 T; 0 other;

## Alignment Scores:

Pred. No.: 5,46e-74 Length: 2817  
 Score: 874.50 Matches: 165  
 Percent Similarity: 77.41% Conservative: 44  
 Best Local Similarity: 61.11% Mismatches: 50  
 Query Match: 59.05% Indels: 11  
 DB: 21 Gaps: 3

US-09-513-365A-1 (1-280) x AAA53790 (1-2817)

QY 18 GlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37

Db 865 GGGGAGCAGATCCACCTCGTGAAC---TATGGGAGGATTACTCGAGTCAATCGAGTCA 921

QY 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln 57

Db 922 CTGCTTTCGACCTCGCAGAGGACGTCCTCGCTGATCGGGAGATCGACGCCAATACCAA 981

QY 58 GluThrLeuLysGluIleAspAspValTyrGluLysTyrLysLysGluAspLeuAsn 77

Db 982 GAGATCCTCGAAGGAGCTGGACGCTACTATGAGAAGTTCAACGGGAGACAGACGGCACC 1041

QY 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGly 97

Db 1042 CAGAAAGCCGGGTACTGCACCTGCATCCAGAGGGCCCTGATCCGAGCAGGAGCTAGGC 1101

QY 98 AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln 117

Db 1102 GATGAGAGATCCAGATCGTGATCAGATGTTGGAGCTGGTGAGAACCGCAGCAGACAG 1161

QY 118 MetGluLeuHisSerGlnCysPhe-----GlnAspProAlaGluSerGluArgAla 134

Db 1162 GTGGACAGTCACGTGGAGCTCTTCAAGCACACAGGACATCAGTCACGGCCTGGTGGC 1221

QY 135 SerAspLysAlaLysMetAspSerSerGlnProGluArg----- 147

Db 1222 AGCGCAAGGGCGGCGCAGACAAAGTCGAAGAGTGAAGCCATCACACAGCAGATAGCCG 1281

QY 148 SerSerArgAtgProArgArgGlnArgThrSerGluSerArgAspLeuCysHisMetAla 167

Db 1282 AATAACACAGCGTCCAGAGGGCGGCAACAAATGAGATCGAGAGAACCGCTCGAATAAT 1341

QY 168 AsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSerLysSerAlaLys 187

Db 1342 CACGACCATGATGACATCACCTCAGAACCGCCCAAGGAGAGAAAGAAACCTCAAAG 1401

QY 188 LysLysLysArgSerIysenAlalysGlnGluArgGluAlaSerProvalGluPheAlaIle 207

Db 1402 AAGAAGAAACCGCTCCCAAGGCCCAACACAGAGAGGGAAGCGCTCTCTGCCCATTC 1461

QY 208 AspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGly 227

Db 1462 GACCCCAACAGCCACGACTCTCTGTGCAACAGGTCTCTACCGGGAGATGATCGGC 1521  
 QY 228 CysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyr 247  
 Db 1522 TGTGACACAGCAAGATGCCCATCGAGTGTTCCTCTCTCGTGGGGCTCAACCAT 1581  
 QY 248 LysProLysGlyLysTrpTyrCysProLysCysArgGlyAspAsnGluLysThrMetAsp 267  
 Db 1582 AAACCAAGGGCAAGTGTACTCCCAAGTGCCTGGGGAGAGCGAGAACCATGGAC 1641  
 QY 268 LysSerThrGluLysThrLysLysAspArg 277  
 Db 1642 AAAGCCCTGGAGAGATCCCAAGAAAGAGAGG 1671

## RESULT 7

AAA53792

ID AAA53792 standard; cDNA; 911 BP.

XX

AC AAA53792;

XX

19-DEC-2000 (first entry)

XX

DE Human P37ING1 coding sequence.

XX

p53; tumour; cancer; detection; antibody; hybridisation; exon 1b;  
 INGI1; ing1; p37ING1; p37ING1; oncogene; gene therapy; diagnosis;  
 proliferation disorder; transformation; transformed cell; human; ds.

XX

OS Homo sapiens.

XX

PN W0200046370-A1.

XX

PD 10-AUG-2000.

XX

PF 04-FEB-2000; 2000WO-US02959.

XX

PR 04-FEB-1999; 99US-0118941.

XX

PA (UNII ) UNIV ILLINOIS FOUND.

XX

PI Gudkov A, Zeremski M, Gurova KV, Grigorian IA;

XX

DR WPI; 2000-491278/43.

XX

DR P-PSDB; AAY97244.

XX

Detecting nucleic acid encoding exon 1b of ing1, useful for diagnosing  
 and treating cancer, comprises contacting sample with isolated nucleic  
 acid comprising sequence of exon 1b and detecting hybridized products

XX

PS Disclosure; Page 122-123; 134pp; English.

XX

Mutations in or loss of the p53 gene occur in more than 50% of  
 human tumours and tumour cell lines, but functional inactivation of  
 the p53 pathway occurs in a much larger proportion of tumours. In  
 many cases the mechanism of functional inactivation of the p53 gene  
 remains unknown but p53 has been found to act in cooperation with  
 INGI1. Functional cooperation between INGI1 and p53 suggested that  
 INGI1 encoded a tumour suppressor protein that functioned within the  
 p53 pathway. This data suggested a possible role for INGI1 in head  
 and neck cancers and chromosomal location of the INGI1 placed it  
 within a region that is frequently rearranged in head and neck  
 cancers. Large scale analysis of tumours involving INGI1 has not  
 revealed mutations in INGI1 nor significant variations in its  
 expression suggesting that INGI1 was not a useful gene to study in  
 cancer etiology. However, alternative initiation exons of the ing1  
 gene, each having their own promoter have been discovered.  
 Expression of one promoter (1a) produces a protein identical to  
 INGI1. Expression of a second promoter (1b) produces a protein having  
 an identical C-terminal fragment to INGI1 but an additional 104  
 N-terminal amino acids. The newly discovered protein has been  
 designated p37ING1 (Wild type: p37ING1). p37ING1 has the  
 characteristics of an oncogene. When overexpressed in cells (even

CC those expressing wild type p53) p37ING1 is able to cause  
 CC proliferation or transformation of those cells. Thus detecting a  
 CC nucleic acid encoding exon 1b of ing1 by hybridisation with an  
 CC isolated nucleic acid having the sequence of exon 1b of ing1  
 CC or its antisense sequence can identify individuals expressing the  
 CC oncogenic form of ing1. Novel peptide sequences taken from the 104  
 CC N-terminal peptide of p37ING1 can also be used to raise antibodies  
 CC that can also be used in detection methods for the p37ING1 variant.  
 CC The polypeptides may be useful in gene therapy for treatment of cell  
 CC proliferation disorders, especially cancers and for diagnosing and  
 CC studying cancers.

XX SQ Sequence 911 BP; 228 A; 255 C; 305 G; 123 T; 0 other;

## Alignment Scores:

Pred. No.:	1-29e-72	Length:	911
Score:	853.50	Matches:	161
Percent Similarity:	75.74%	Conservative:	45
Best Local Similarity:	59.19%	Mismatches:	51
Query Match:	57.63%	Indels:	15
DB:	21	Gaps:	3

US-09-513-365A-1 (1-280) x AAA53792 (1-911)

QY 18 GlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37  
 Db 63 GGGGAGCAGCTCCACCTGGTGAAC---TATGTGGAGGACTACTGTGACTCCATCGAGTCC 119  
 QY 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln 57  
 Db 120 CTGCCTTTCGACTTGCAGAGAAATGTCTCGCTGATCGGGAGATCGACGCGAAATACCA 179  
 QY 58 GluThrLeuLysGluIleAspValTyrGluLysTyrLysGluAspLeuAsn 77  
 Db 180 GAGATCCTGAAGAGCTGACAGGTGTCTACGAGCGCTTCAGTCCGAGAGACAGACGGGCG 239  
 QY 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLysGly 97  
 Db 240 CAGAGCGCGGATGCTGCTGTCAGCGCGCGCTGATCCGACGCCAGGAGCTGGGC 299  
 QY 98 AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln 117  
 Db 300 GACGAGAGATCCAGATCGTGCAGCAGATGTGTGAGTGTGTGAGAACCGCGCGCAG 359  
 QY 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
 Db 360 GTGGACAGCCAGCTGGAGCTGTTCGAGCGCGCAGCAGGAGCTGGCGCACACAGTGGGCAAC 419  
 QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnPro 145  
 Db 420 AGCGGCAGGTGGCGCGCAGCAGGCCAATGGCGGATGCGTAGCGGAGCTGTGACAGCCC 479  
 QY 146 GluArgSerArgArgProArgGlnArgThrSerGluSerArgAspLeuCysHis 165  
 Db 480 -----AACAGCAAGCGCTCACCGCGGCGCGCAACACAGAGAACCGTGAGACCGCTCC 533  
 QY 166 MetAlaAsnGlyIleGluAspCysAspGlnProProLysGluLysLysSerLysSer 185  
 Db 534 AGCAACACAGCAGCAGCAGCGCGCTCCGCGGCACACCAAGAGAGAGAGGCAAGAC 593  
 QY 186 AlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPhe 205  
 Db 594 TCAGAGAGAGAGAGCGCTCCAGGCCAAGCGGAGCGGAGAGCGGTCCCTCGCCGACCTC 653  
 QY 206 AlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMet 225  
 Db 654 CCCATCGACCCCAAGCAACCCACGACTGTCTGTGCAACCAAGGTCTCTATGGGAGATG 713  
 QY 226 IleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeu 245  
 Db 714 ATCGGCTGCGACACAGCAGTGCCTCCATCGAGTGTGTTCCTCTCTCGTGGTGGGCTC 773  
 QY 246 ThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspAsnGluLysThr 265







PA (UNII ) UNIV ILLINOIS FOUND.  
 XX  
 PI Gudkov, A., Zeremski M., Gurova KV, Grigorian IA;  
 XX  
 DR WPI; 2000-491278/43.  
 DR  
 XX  
 PT Detecting nucleic acid encoding exon 1b of ing1, useful for diagnosing  
 PT and treating cancer, comprises contacting sample with isolated nucleic  
 PT acid comprising sequence of exon 1b and detecting hybridized products  
 XX  
 PS Claim 17; Fig 6; 134pp; English.  
 XX  
 CC Mutations in or loss of the p53 gene occur in more than 50% of  
 CC human tumours and tumour cell lines, but functional inactivation of  
 CC the p53 pathway occurs in a much larger proportion of tumours. In  
 CC many cases the mechanism of functional inactivation of the p53 gene  
 CC remains unknown but p53 has been found to act in cooperation with  
 CC ING1. Functional cooperation between ING1 and p53 suggested that  
 CC ING1 encoded a tumour suppressor protein that functioned within the  
 CC p53 pathway. This data suggested a possible role for ING1 in head  
 CC and neck cancers and chromosomal location of the ING1 placed it  
 CC within a region that is frequently rearranged in head and neck  
 CC cancers. Large scale analysis of tumours involving ING1 has not  
 CC revealed mutations in ING1 nor significant variations in its  
 CC expression suggesting that ING1 was not a useful gene to study in  
 CC cancer etiology. However, alternative initiation exons of the ing1  
 CC gene, each having their own promoter have been discovered.  
 CC Expression of one promoter (1a) produces a protein identical to  
 CC ING1. Expression of a second promoter (1b) produces a protein having  
 CC an identical C-terminal fragment to ING1 but an additional 104  
 CC N-terminal amino acids. The newly discovered protein has been  
 CC designated p37ING1 (wild type: p33ING1). p37ING1 has the  
 CC characteristics of an oncogene. When overexpressed in cells (even  
 CC those expressing wild type p53) p37ING1 is able to cause  
 CC proliferation or transformation of those cells. Thus detecting a  
 CC nucleic acid encoding exon 1b of ing1 by hybridisation with an  
 CC isolated nucleic acid having the sequence of exon 1b of ing1  
 CC or its antisense sequence can identify individuals expressing the  
 CC oncogenic form of ing1. Novel peptide sequences taken from the 104  
 CC N-terminal peptide of p37ING1 can also be used to raise antibodies  
 CC that can also be used in detection methods for the p37ING1 variant.  
 CC The polypeptides may be useful in gene therapy for treatment of cell  
 CC proliferation disorders, especially cancers and for diagnosing and  
 CC studying cancers.  
 XX  
 SQ Sequence 1835 BP; 541 A; 387 C; 468 G; 439 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8,55e-62 Length: 1835  
 Score: 745.00 Matches: 140  
 Percent Similarity: 76.75% Conservative: 35  
 Best Local Similarity: 61.40% Mismatches: 43  
 Query Match: 50.30% Indels: 10  
 DB: 21 Gaps: 2  
 US-09-513-365A-1 (1-280) x AA453789 (1-1835)  
 QY 60 LeuLysGluLeuAspValTyrGluLysTyrLysGluAspLeuAsnGlnLys 79  
 Db 6 CTGAAGGAGCTGCGACGACTACTATGAGAGTTCAACGGGAGAGACAGCGCAGAGAG 65  
 QY 80 LysArgLeuGlnGlnLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGlyAspGlu 99  
 Db 66 CGCGGGGTACTGCACTGCTCCAGAGGGCCCTGATCCGACGAGGAGCTAGCGATGAG 125  
 QY 100 LysileGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGlu 119  
 Db 126 AAGATCCAGATCTGTGATGAGTGGTGGAGCTGGTGGAGAACCCGACGACAGGTGGAC 185  
 QY 120 LeuHisSerGlnCysPhe-----GlnAspProAlaGluSerGluArgAlaSerAsp 136  
 Db 186 AGTCAGGTGGAGCTCTTCGAAGCACACCGACGACATCATGTCACGGCTGCTGGCAGCGGC 245

QY 137 LysAlaLysMetAspSerSerGlnProGluArg-----SerSer 149  
 Db 246 AAGGCGGGCCAGGACAGTTCGAAGATGAGGCCCATCACACGCGACAGATAAGCGAATAAC 305  
 QY 150 ArgArgProArgArgGlnArgThrSerGluSerArgAspLeuCysHisMetAlaenGly 169  
 Db 306 AAGCGGTCCAGGAGGCGAGCAAAATGAGATTCGAGAGAACGCGTCGAATAATCATCAGAC 365  
 QY 170 IleGluAspCysAspAspGlnProProLysGluLysLysSerLysSerAlaLysLysLys 189  
 Db 366 CATGATCATCATCCTCAGGAACGCCAAGGAGAGAAAGCAAAACCTCAAGAAGAAG 425  
 QY 190 LysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIleAspPro 209  
 Db 426 AAACGCTCCAGGCCAAACGACGAGAGGAGAGGCTCTCTCCGCGACCTTCCCATCGACCCC 485  
 QY 210 AsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAsp 229  
 Db 486 AAGGAGGCCACGCTACTGTCTGTGCAACAGGTCTCTACGGGGAGATGATCGCTGTGAC 545  
 QY 230 AsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyrLysPro 249  
 Db 546 AACGCAAGTATGCTGCTGCTCCCAAGTCCGCTGGGGAGCGAGAACCATGACCAAGGCC 605  
 QY 250 LysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGluLysThrMetAspLysSer 269  
 Db 606 AAGGCAAGTGTGCTGCTGCTCCCAAGTCCGCTGGGGAGCGAGAACCATGACCAAGGCC 665  
 QY 270 ThrGluLysThrLysLysAspArg 277  
 Db 666 CTGGAGAGTCCCAAGAAAGAGAGG 689  
 RESULT 11  
 AAH28479  
 ID AAH28479 standard; DNA; 1143 BP.  
 AC AAH28479;  
 XX  
 DT 17-SEP-2001 (first entry)  
 XX  
 DE Nucleotide sequence of a human cancer associated antigen.  
 XX  
 KW Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 16..900  
 FT /tag= a  
 FT /transl\_except= "(pos: 25..30, aa: Cys)"  
 FT /transl\_except= "(pos: 124..126, aa: Pro, Ala)"  
 FT /product= "cancer associated antigen"  
 XX  
 PN WO200147959-A2.  
 XX  
 PD 05-JUL-2001.  
 XX  
 PF 29-NOV-2000; 2000WO-US42334.  
 XX  
 PR 30-NOV-1999; 99US-0451739.  
 PR 24-OCT-2000; 2000US-0602362.  
 XX  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 PA (SLOK) SLOAN KETTERING INST CANCER RES.  
 PA (CORR) CORNELL RES FOUND INC.  
 XX  
 XX Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;  
 XX WPI; 2001-441706/47.  
 DR P-PSDB; AAB84697.  
 XX  
 PT Isolated cancer associated nucleic acid molecule identified by SPREX  
 PT (serological identification of antigens by recombinant expression



QY 96 LeuGlyAspGluLeuIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
 DB 151 CTGGCGGACGAGAAATCAGATCTGGAGCAGATGGTGGAGAACCGCAGC 210  
 QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
 DB 211 CGGAGGTGGACGACCGATGCTGTCGAGCGCAGCAGGAGCTGGCGCACACAGTG 270  
 QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
 DB 271 GGCAACAGCGGCAAGGTTGGCGGACAGCCCAATGGCGATGCGTAGCGCAGTCTGAC 330  
 QY 144 GlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSerArgAspLeu 163  
 DB 331 AAGCCCC-----AACAGCAAGCGCTCACGGCGGCGAGCAACACAGAAACCGTGAGAAC 384  
 QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSer 183  
 DB 395 CGGTCGACCAACACGACACGACGCGGCGCTCGGGCACACCCACAGGAGAGAGGCC 444  
 QY 184 LysSerAlaLysLysLysLysArgSerLysAlaLysGlnLysGlnLysGlnLysProVal 203  
 DB 445 AAGACCTCCAGAGAGAGAGCGCTCCAGAGCGGAGCGGAGAGCGCTCCCTGCC 504  
 QY 204 GluPheAlaIleAspProAsnGluProThrCysLeuCysAsnGlnValSerTyrGly 223  
 DB 505 GACCTCCCATCGACCCCAACACCCAGCTACTGTCTGTGCAACAGAGTCTCTATGG 564  
 QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243  
 DB 565 GAGATGATCGGTGGACACACAGAGTCCCATCGAGTGTTCATCTCTCGTGGTG 624  
 QY 244 SerLeuThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspGln 263  
 DB 625 GGGCTCAATCATAAACCCCAAGGCAAGTGTACTGTCTCCCAAGTGGCGGGGAGAACGAG 684  
 QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
 DB 685 AAGACCATGGACAAAGCCCTGGAGAAATCCAAAAAGAGAGG 726  
 RESULT 13  
 AAV62285  
 ID AAV62285 standard; cDNA; 1902 BP.  
 XX  
 AC AAV62285;  
 DT 18-JAN-1999 (first entry)  
 DE Partial INGI partial cDNA sequence.  
 KW INGI gene; p33ING1; human; apoptosis; cell death; breast cancer;  
 KW brain tumour; gene therapy; tumour suppressor; ss.  
 OS Homo sapiens.  
 FX  
 FT Key Location/Qualifiers  
 FT CDS 109..741  
 FT /\*tag= a  
 XX  
 PN W09844102-A2.  
 XX  
 PD 08-OCT-1998.  
 XX  
 PF 26-MAR-1998; 98WO-CA00277.  
 XX  
 PR 27-MAR-1997; 97US-0828158.  
 XX  
 PA (UYTE-) UNIV TECHNOLOGIES INT INC.  
 XX  
 PI Garkavtsev I, Helbing CC, Johnston RN, Riabowol K;  
 XX WPI; 1998-542700/46.  
 DR

P-PSDB; AAW79674.

Modulating eukaryotic apoptosis by increasing p33ING1 activity -  
 using p33ING1 derivatives, to induce apoptosis in cancer cells, and  
 in the investigation of apoptotic pathways

Example 2; Fig 2; 66pp; English.

This is the nucleotide sequence of a human INGI (inhibitor of  
 growth) partial cDNA clone that codes for a p33ING1 polypeptide  
 (see AAW79674), a novel inhibitor of cell growth and a candidate  
 tumour suppressor. INGI is a new gene that is expressed in normal  
 mammary epithelial cells, but which is expressed only at lower  
 levels in several cancerous mammary epithelial cell lines and is  
 not expressed in many primary brain tumours. To isolate INGI, a  
 subtractive hybridisation of breast cancer cell line cDNAs was  
 performed with cDNA from normal mammary epithelial cells, and  
 subtracted cDNAs were cloned into retrovirus vector pLNCX.  
 Following passage through a packaging line, normal mouse mammary  
 epithelial cells were infected, and infected cells were injected  
 into nude mice. Putative transforming fragments from tumours were  
 isolated by PCR (see AAV62290-91) and subcloned into INCX. An INGI  
 fragment was obtained and used to screen normal human fibroblast  
 and HeLa cell cDNA libraries. 2 Clones were sequenced to obtain  
 the partial INGI sequence. The complete cDNA sequence (see  
 AAV62292) was obtained by RACE. A claimed method to potentiate  
 apoptosis in a eukaryotic cell involves administering an active  
 p33ING1 peptide or an oligonucleotide encoding such a peptide.  
 A claimed method for inhibiting apoptosis in a eukaryotic cell  
 involves administering an antisense oligonucleotide. Also claimed  
 are a method for determining the apoptotic characteristics of a  
 eukaryotic cell, an assay for determining the level of p33ING1  
 activity in a eukaryotic cell, and an isolated eukaryotic cell  
 substantially free of p33ING1 biological activity. The invention  
 discloses INGI derivatives or variants that may be used to induce  
 apoptosis in eukaryotic cancer cells.

Sequence 1902 BP; 574 A; 390 C; 462 G; 476 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 3.85e-60 Length: 1902  
 Score: 728.00 Matches: 136  
 Percent Similarity: 74.79% Conservative: 39  
 Best Local Similarity: 58.12% Mismatches: 45  
 Query Match: 49.16% Indels: 14  
 DB: 19 Gaps: 2

US-09-513-365A-1 (1-280) x AAV62285 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspArg 75  
 DB 31 TGGAAACAGATCTCGAAGGAGCTAGACGAGTGTACGAGCGCTTCAGTCGCGAGACAGAC 90  
 QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlu 95  
 DB 91 GGGCGCAGAGCGCGGATGCTGCATGTGTGACGCGCGCTGATCCGACGACGAGAG 150  
 QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
 DB 151 CTGGCGGACGAGAAATCAGATCTGAGCAGATGTTGGAGTGTGTGGAGAACCGCAGC 210  
 QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
 DB 211 CGGAGGTGGACGACCGATGCTGTCGAGCGCAGCAGGAGCTGGCGCACACAGTG 270  
 QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
 DB 271 GGCAACAGCGGCAAGGTTGGCGGACAGCCCAATGGCGATGCGTAGCGCAGTCTGAC 330  
 QY 144 GlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSerArgAspLeu 163  
 DB 331 AAGCCCC-----AACAGCAAGCGCTCACGGCGGCGAGCAACACAGAAACCGTGAGAAC 384

QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspGlnProProLysGluLysLysSer 183  
 Db 385 GCGTCCAGCAACACGACACGACGCGCGCTCGGGCACACCCCAAGGAGAAAGGCC 444  
 QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
 Db 445 AAGACTTCCAAAGAAAGACGCTCCAAAGCCAAAGCGGAGGAGCGCTCCCTGCC 504  
 QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
 Db 505 GACTCTCCCATGACCCCAACGACACCATGCTGTGTGCAACAGGTCTCTATGGG 564  
 QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGlnTrpPheHisPheSerCysVal 243  
 Db 565 GAGATGATCGCTCGCACACGACGAGTGCCTCATCGAGTGGTTCACCTTCTCGTGGTG 624  
 QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263  
 Db 625 GGGCTCAATCAATAAACCAAGGGCAAGTGTGTCTCCCAAGTCCCGGGGGAGAACGAG 684  
 QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
 Db 685 AAGACCATGACCAAGCCCTGGAGAAATCCAAAAGAGAGAGG 726  
 RESULT 14  
 AAD46126  
 ID AAD46126 standard; cDNA; 1905 BP.  
 XX  
 AC AAD46126;  
 XX  
 DT 27-DEC-2002 (first entry)  
 XX  
 DE Human tumour suppressor protein encoding cDNA.  
 XX  
 KW Human; tumour suppressor protein; cell proliferative disorder; vaccine;  
 KW inflammation; brain cancer; adenocarcinoma; cervix cancer; bone cancer;  
 KW apoptosis; leukaemia; lymphoma; melanoma; therapy; chromosome 13; gene;  
 KW ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..160  
 FT FT /\*tag= a  
 FT CDS 161..1036  
 FT FT /\*tag= b  
 FT FT /product= "Human tumour suppressor protein"  
 FT 3'UTR 1037..1905  
 FT FT /\*tag= c  
 XX  
 PN W0200268468-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 05-FEB-2002; 2002WO-US03235.  
 XX  
 PR 27-FEB-2001; 2001US-0793706.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Gong F, Yan C;  
 XX  
 DR WPI; 2002-698658/75.  
 DR P-PSDB; AAE28633.  
 XX  
 PT New human tumor suppressor proteins, useful for developing human  
 PT therapeutic agents, or preventing or treating inflammation, or  
 PT disorders associated with cell proliferation, e.g. bone cancer, brain  
 PT cancer, leukemia or lymphoma  
 XX  
 PS Claim 4; Fig 1; 200pp; English.  
 XX  
 CC The invention relates to human tumour suppressor polypeptides and

CC polynucleotides. Sequences of the invention are useful for diagnosing,  
 CC preventing or treating inflammation, or disorders associated with cell  
 CC proliferation and apoptosis e.g. bone cancer, brain cancer, cervix  
 CC cancer, adenocarcinoma, leukaemia, lymphoma or melanoma. They are  
 CC particularly useful as models for developing human therapeutic targets,  
 CC identifying therapeutic proteins, or serving as targets for the  
 CC development of human therapeutic agents that modulate tumour suppressor  
 CC protein activity in cells and tissues that express the tumour suppressor  
 CC protein. Polypeptides of the invention are used for identifying agents  
 CC that modulate their activity. They are useful for raising antibodies or  
 CC eliciting an immune response; as a reagent in assays designed to  
 CC quantitatively determine levels of the protein (or its binding partner  
 CC or ligand) in biological fluids; or as markers for tissues in which the  
 CC corresponding protein is preferentially expressed. The invention is also  
 CC used as vaccines. The present sequence is human tumour suppressor  
 CC protein encoding cDNA. The tumour suppressor gene is located on  
 CC chromosome 13.  
 XX

SQ Sequence 1905 BP; 532 A; 428 C; 528 G; 417 T; 0 other;

#### Alignment Scores:

Pred. No.: 3,868-60 Length: 1905  
 Score: 728.00 Matches: 136  
 Percent Similarity: 75.22% Conservative: 37  
 Best Local Similarity: 59.13% Mismatches: 47  
 Query Match: 49.16% Indels: 10  
 DB: 24 Gaps: 2

US-09-513-365A-1 (1-280) x AAD46126 (1-1905)

QY 58 GluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspLeuAsn 77  
 Db 332 GAGATCTTGAAGAGAGTAGACGAGTGTACGAGCGCTTCAGTCGCAGACAGCGGGCG 391  
 QY 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGly 97  
 Db 392 CAGAAGCGCGGATGTGCTGTGTGACGCGCGCTGATCCGACGAGGAGCTGGGC 451  
 QY 98 AspGluLysIleGlnIleValThrGlnMetLeuGluValClnAsnArgAlaArgGln 117  
 Db 452 GACGAGAAGATCCAGATCGTGAGCCAGATGGTGAGAGACCGCGACGCGCGCAG 511  
 QY 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGlu---SerGluArgAlaSerAsp 136  
 Db 512 GTGACAGCCACGTGGAGCTGTTGAGGCGCAGCAGGAGCTGGCGCACAGCGGCAAC 571  
 QY 137 LysAlaLysMetAspSerSerGlnPro-----GluArg 147  
 Db 572 AGCGGCAGGCTGGCGCGACAGCGCCAAAGCGAGCGGCGAGCGGCTGACAAAGCCC 631  
 QY 148 SerSerArgArgProArgArgGlnArgThrSerGluSerArgAspLeuCysHisMetAla 167  
 Db 632 AACAGCAGCGCTCACGGCGGCGAGCGCAACACAGAGAACCGTGAGAACCGCGTCCAGCAAC 691  
 QY 168 AsnGlyIleGluAspCysAspGlnProProLysGluLysLysSerLysSerAlaLys 187  
 Db 692 CAGCACACGACGACGGCGCTCGGGCACACCCAGGAGAGAGAGCGGAGAGCTCCCAAG 751  
 QY 188 LysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIle 207  
 Db 752 AAGAAGAGCGCTCCAAAGGCCAAGCGGAGCGAGAGGCGTCCCTGCGCGACCTCCCATC 811  
 QY 208 AspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGly 227  
 Db 812 GACCCCAACGAAACCCACGTACTGTCTGTGCAACACCGAGTCTCTATGGGAGATGATCGGC 871  
 QY 228 CysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyr 247  
 Db 872 TCGACAAACGACGAGTGGTCCCATCGAGTGTTCACCTTCTCGTGGCGGCTCAATCAT 931  
 QY 248 LysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGluLysThrMetAsp 267  
 Db 932 AAMCCCAAGGCAAGTGTGTACTGTCTCCCAAGTCCCGGGGGAGAGACGAGAACCATGGAC 991



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2004, 19:38:43 ; Search time 2031 Seconds  
(without alignments)  
3350.691 Million cell updates/sec

Title: US-09-513-365A-1  
Perfect score: 1481  
Sequence: 1 MLGQQQQQLYSSAALLTGER.....DNEKTMDSKTEKKDRSR 280

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlb  
-Q=/cgr2\_1/USPTO.spool\_p/US09513365/runat\_12012004\_163932\_28778/app\_query.fasta\_1.455  
-DB=EST -Qfmt=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513365 @CGN 1 1 3549 @runat\_12012004\_163932\_28778 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estom:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C	1	1433.5	96.8	2643	11	AK048800 Mus muscu
	2	1433.5	96.8	2812	11	AK083144 Mus muscu
	3	1395.5	94.2	1464	11	AK012716 Mus muscu
	4	1296	87.5	912	13	BQ277444 AGENCOURT
	5	1123	75.8	736	13	BQ225297 603799281
	6	1121	75.7	892	13	BQ332089 603869689
	7	1096.5	74.0	919	10	BG211544 RST31111
	8	1094	73.9	793	10	BE796780 601587557
	9	1063	71.8	778	12	BI548536 603191255
	10	1043	70.4	779	13	BU259874 603504254
C	11	1027	69.3	673	13	BU386788 603581891
	12	1024	69.1	735	12	BM982877 UI-CF-ENI
	13	1022	69.0	707	13	BU323498 603489003
	14	1007.5	68.0	864	13	BU591073 AGENCOURT
C	15	1006.5	68.0	798	10	BG184056 RST2972 A
	16	965	65.2	927	14	CA973890 AGENCOURT
C	17	961	64.9	664	12	BM719460 UI-E-EJ1-
	18	950	64.1	728	13	BU613909 UI-M-FRO-
	19	935	63.1	708	12	BI460319 603201967
	20	933.5	63.0	847	13	BU196949 AGENCOURT
C	21	920	62.1	675	12	BM682789 UI-E-EJ1-
	22	914	61.7	671	13	BU404104 604138228
C	23	911.5	61.5	859	13	BU900292 AGENCOURT
	24	902	60.9	891	12	BI091047 602855235
C	25	897	60.6	541	10	BF114684 7J66C08.X
	26	889	60.0	679	9	AV647917 AV647917
C	27	883	59.6	564	10	BE839427 RC3-PN014
	28	874	59.0	538	10	BF523624 UI-R-C1-1
C	29	864	58.3	653	12	BI495873 df121d04.
	30	858	57.9	671	9	AL857736 AL857736
C	31	856	57.8	522	12	BI495874 df121d04.
	32	852	57.5	566	10	BE839460 RC3-PN014
C	33	851	57.5	651	12	BP110430 BP110430
	34	849	57.3	720	9	AA797495 vW28001.X
C	35	847	57.2	703	12	BI546850 603189754
	36	830	56.0	531	9	AA204672 z187C06.X
C	37	829	56.0	525	9	AI885642 w159h01.X
	38	828	55.9	593	13	BU264721 603814832
C	39	816	55.1	869	13	BU226578 603796130
	40	814	55.0	616	9	AL673816 AL673816
C	41	814	55.0	692	14	CB937790 IPCGX13
	42	809	54.6	941	14	CA982474 AGENCOURT
C	43	805	54.4	547	9	AI186701 q882d12.X
	44	798	53.9	482	10	BE831862 RC6-MT006
45	794	53.6	648	9	AL874284 AL874284	

ALIGNMENTS

RESULT 1  
AK048800  
LOCUS  
DEFINITION  
Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:C230071A14 product:inhibitor of growth  
family, member 1-like, full insert sequence.  
ACCESSION  
AK048800  
VERSION  
AK048800.1 GI:26339579  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Muridae; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999) **MB**  
 MEDLINE 99279253  
 PUBMED 10349636

## REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20493374  
 PUBMED 11042159

## REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

## REFERENCE

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gaasterland, T., Giess, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
 Quackenbush, J., Schram, L. M., Stauble, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Saya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
 Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,  
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.  
 and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 MEDLINE 21085660  
 PUBMED 11217851

## REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 MEDLINE 12606170  
 PUBMED 12026433

## REFERENCE

6 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yaeunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission

## REFERENCE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome

## COMMENT

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

Location/Qualifiers  
 1..2643  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="PANTOM DB:C230071A14"  
 /db\_xref="taxon:10090"  
 /clone="C230071A14"  
 /cissue\_type="cerebellum"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="0 day neonate"  
 476..1321  
 /note="unnamed protein product; inhibitor of growth  
 family, member 1-like (MGD|MG1:1916510, GB|NM\_023503,  
 evidence: BLASTN, 98%, match=971)  
 putative"

## CDS

683 a 557 c 710 g 693 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,338-82 Length: 2643  
 Score: 1433.50 Matches: 271  
 Percent Similarity: 98.58% Conservative: 6  
 Best Local Similarity: 96.44% Mismatches: 3  
 Query Match: 96.79% Indels: 1  
 DB: 11 Gaps: 1

## BASE COUNT

US-09-513-365A-1 (1-280) x AK048800 (1-2643)  
 QY 1 MetLeuGly---GlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGlu 19  
 Db 476 ATGTTAGGCGCAGCAGCAGCAGCAGCTGTACTCGCGCGCGCTCTCGCCGAGAG 535  
 QY 20 ArgSerArgLeuLeuThrCystTyrValGlnAspTyrLeuGluCysValGluSerLeuPro 39  
 Db 536 CGAGCGCGCT 595  
 QY 40 HisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThr 59  
 Db 596 CACGACATGCAGAGGACGCTGTCTCGTCTCGGAGCTGGACACAAATACCAAGAACG 655  
 QY 60 LeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLys 79  
 Db 656 TTAAGGAAATTCATGATGCTCTATCAAAATATAAGAAAGAGATGATTCACCAACAGAAA 715  
 QY 80 LysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGlyAspGlu 99  
 Db 716 AAACCCCTACGACGATCTCCAGAGAGCGTTAATCAATAGCAAGAAATGGAGATGAA 775  
 QY 100 LysileGlnileValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGlu 119



Db 776 AAAATTTCAGATTGTCACCCAGATCTCGAATTCGGTGGAGAACCGACGAGACAAATGGAG 835

QY 120 LeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLys 139

Db 836 CTGCAATTCAGTGTTCAGAGTCTGCTGAAAGTGAGCGACCTCAGACAAAGTCGAAG 895

QY 140 MetAspSerSerGlnProGluArgSerArgProArgProArgGlnArgThrSerGlu 159

Db 896 ATGATTTCCAGTCAACCGGAAGATCTCTAGAGACCTCTGAAGACAGAGGACCACTGAG 955

QY 160 SerArgAspLeuGlyCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLys 179

Db 956 AGCGGTGATTTATGTACATGACAAACGGATTCAGACCTGTGATGATCAACACCGAAA 1015

QY 180 GluLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLys 199

Db 1016 GAAAGAGATCCAGTCCGCAAGAGAGAACCGCTCCAGGCGCAAGCAGGAGGGAG 1075

QY 200 AlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuGlnGlu 219

Db 1076 GCATCCCTGTCAGTGTTCATGATCCCAATGAGCCACCTACTGTGTGTAACCAA 1135

QY 220 ValSerTyrGluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrPheHis 239

Db 1136 GTCTCTACGGGAGATGATAGCTGTGCAATGAACAGTGTCCCATGAATGGTTTCAC 1195

QY 240 PheSerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArg 259

Db 1196 TTTTCATGTGTTTCATCTACCTATAACCAAGGAGGATGTTGTTGCCAAAGTGTAGG 1255

QY 260 GlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSer 279

Db 1256 GGAGACAATCAGAAACCAATGAGTGGCAAAAGTACCGAAAAGACAAAGGAGAGAGCG 1315

QY 280 Arg 280

Db 1316 AGG 1318

RESULT 2

AK083144

LOCUS

DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone: C630016P10 product: inhibitor of growth family, member 1-like, full insert sequence.

ACCESSION AK083144

VERSION 1

KEYWORDS GI:26350286

SOURCE HTC; CAP trapper.

ORGANISM Mus musculus (house mouse)

REFERENCE 1 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Washiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system -384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustinich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001).

21085660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2812)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

TITLE

JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1. .2812  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:C630016P10"  
/db\_xref="taxon:10090"  
/clone="C630016P10"  
/sex="male"

/tissue type="hippocampus"  
/clone lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
479..1324  
CDS

/notes="unnamed protein product; inhibitor of growth  
family, member 1-like (MGJ|MGI:1916510, GB|NM\_023503,  
evidence: BLASTN, 98%, match=971)  
putative"  
/codon\_start=1  
/protein\_id="BAC38783.1"  
/db\_xref="GI:26350287"  
/translation="MLGQQQQOOLYSSAALLTGERSLRLSCYVODYLECVESLPHDMQ  
RNVSLRELDNRYQETLEIDVYKYKKEDDSNKKRLQQLORALINSQELGDEKI  
QIVTQMLEVNRAROMELHSOCFODPAESRASDKMSQOPSSRRPRPRTISE  
SDLCMTNGIDDCDDPPKXRSKAKKXSKAKQEREASFPVPAIDNPFTYCLC  
NOVSGEMIGCDNEQCPIEWFHSVSLTYKPKRWYCPKCRGDNEKTMKRSKTEKTK  
ERROR"

BASE COUNT 725 a 589 c 743 g 753 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 1.39e-82 Length: 2812  
Score: 1433.50 Matches: 271  
Percent Similarity: 98.58% Conservative: 6  
Best Local Similarity: 96.44% Mismatches: 3  
Query Match: 96.79% Indels: 1  
DB: 11 Gaps: 1

US-09-513-365A-1 (1-280) x AK083144 (1-2812)

QY 1 MetLeuGly---GlnGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGlu 19  
Db 479 ATGTTAGGAGCAGCAGCAGCAGCAGCTGTACTCTGCGCCCGCTCTCTGACCGGAGAG 538  
QY 20 ArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuPro 39  
Db 539 CGGAGCGCGTCT 598  
QY 40 HisAspMetGlnArgSerValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThr 59  
Db 599 CACGACATGACAGAGAACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658  
QY 60 LeuLysGluLeuLeuAspValTyrGluLysTyrLysLysGluAspLeuAsnGlnLys 79  
Db 659 TTAAGGAATTCATGATGCTATGATAAATAATAGAAAGAGATGATTCACACACAGAAA 718  
QY 80 LysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlnLeuGlyAspGlu 99  
Db 719 AAACGGCTACAGCAGCATCTCCAGAGAGCGTTAATCAATAGCCCAAGATTTGGGAGATGAA 778  
QY 100 LysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGlu 119  
Db 779 AAAATTTCAGATTCTACCCAGATGCTCGAATTGCTGGAGAACCCGAGCAGACAAATGGAG 838  
QY 120 LeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLys 139  
Db 839 CTCGATTTCAGGTGTTTCAGGATCTCTCTGAAAGTGAGCGACCTCAGACAAAGTGAAG 898  
QY 140 MetAspSerSerGlnProGluArgSerSerArgProArgArgGlnArgThrSerGlu 159  
Db 899 ATGGATTCCAGTCAACCGGAAGATCTCTAGAGACCTCGAAGACAGAGACCACTGAG 958  
QY 160 SerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLys 179  
Db 959 AGCCGTGACTTATGTACATGACAAACCGGATTGACGACTGTGATCATCAACCCAGAAA 1018  
QY 180 GluLysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGlu 199  
Db 1019 GAAAGAGATCCCAAGTCCCGCAAGAGAAAGCGCTCCCAAGCCCAAGCAGGAGAGGGAG 1078  
QY 200 AlaSerProValGluPheAlaLysAspProAsnGluProThrTyrCysLeuCysAsnGln 219  
Db 1079 GCATCCCTGTCAGTTTGCATTCGATCCCAATGAGCCCACTTACTTGTGTACCAA 1138

QY 220 ValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrPheHis 239  
Db 1139 GTGTCTTACGGGAGATGATAGCTGTGACATGACAGTGTCCCATGTAATGTTTCAC 1198  
QY 240 PheSerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArg 259  
Db 1199 TTTTCATGTGTTTCACTCACTTATTAACCCCAAGGGGAAATGATTTGCCAAAGTGTAGG 1258  
QY 260 GlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSer 279  
Db 1259 CGAGCAATGAGAAAACCATGACAAAAGTACCGAAAAGACAAAAGAGAGAGCGC 1318  
QY 280 ATG 280  
Db 1319 AGG 1321  
RESULT 3  
AK012716 1464 bp mRNA linear HTC 05-DEC-2002  
LOCUS Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
DEFINITION enriched library, clone:281001M06 product:inhibitor of growth  
family, member 1-like, full insert sequence.  
ACCESSION AK012716  
VERSION AK012716.1 GI:12849649  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Mateu, Y., Nakaide, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staib, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarrelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

**TITLE**  
Functional annotation of a full-length mouse cDNA collection  
**JOURNAL**  
Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

RESULT 4  
 BQ277444 912 bp mRNA linear EST 07-MAY-2002  
 LOCUS AGENCOURT\_6738254 NIH\_MGC\_127 Homo sapiens cDNA clone IMAGE:5810745  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ277444  
 VERSION BQ277444.1 GI:20487652  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 912)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: NCI  
 CDNA Library Preparation: Michael Brownstein Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1CM2058 row: h column: 10  
 High quality sequence stop: 587.  
 Location/Qualifiers

## FEATURES

1..912  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5810745"  
 /tissue\_type="mixed (pool of 40 RNAs)"  
 /lab\_host="PH10B (T1-phage-resistant)"  
 /clone\_lib="NIH MGC 127"  
 /notes="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);  
 Site 2: SfiI (ggccctcgcc); Double-stranded cDNA was  
 prepared from a pool of 40 cell line polyA+ RNAs (bladder  
 - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
 kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
 ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
 gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
 used in cloning as follows:  
 5'-AACGATGGTATCAGCGAGTGGCCATTACGGCGGG-3' and  
 5'-ATTCTAGAGCGGAGCGGCGGACATG-3'. Full-length  
 enriched library was constructed using the Clontech  
 Creator SMART kit and size-selected to contain the 1-2 kb  
 size fraction (other fractions present in NIH MGC 126 and  
 NIH MGC 128). Library created in the laboratory of T.  
 Udén, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC  
 Library."

BASE COUNT 313 a 184 c 235 g 180 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,04e-74 Length: 912  
 Score: 1296.00 Matches: 261  
 Percent Similarity: 92.71% Conservativeness: 6  
 Best Local Similarity: 90.62% Mismatches: 11  
 Query Match: 87.51% Indels: 11  
 DB: 13 Gaps: 1

US-09-513-365A-1 (1-280) x BQ277444 (1-912)

QY 1 MetLeuGlnGlnGlnGlnGlnLeuTyrSerSerAlaLeuLeuThrGlyGluArg 20  
 Db 22 AUGTTAGGCGAGCAGCAGCAGCAACTGACTCTCGCTCGCTCTCTGACCGGGAGCGG 81  
 QY 21 SerArgLeuLeuThrCystyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40  
 Db 82 AGCCGGCTGCTCACCTGTACGTGCGAGGACTACCTTGAGTGGGTGGGTCGCTGCCAC 141

QY 41 RepMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60  
 Db 142 GACATGCGAGGAAAGCTGTCTGTGTCGAGAGCTGGCAACAATATCAAGAAAGTTA 201  
 QY 61 LysGluLeuAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLys 80  
 Db 202 AAGGAAATTGATGTCTACGAAATATTAAGAAAGAAGATGATTTAAACCAAGAGAAA 261  
 QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGlyAspGluLys 100  
 Db 262 CGTCTACAGCAGCTTCTCCAGAGACACTAATTAATAGTCAAGAATTGGGAGATGAAAA 321  
 QY 101 IleGlnIleValThrGlnMetLeuGlnValGluAsnArgAlaArgGlnMetGluLeu 120  
 Db 322 ATACAGATTGTACACAATGCTCGAATGGTGGAAATCGGCAAGACAAATGGNGTTA 381  
 QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysLysMet 140  
 Db 382 CACTCACAGTGTTCCTCAAGATCTCTGCTGAAAGTGAACGAGCTCAGATAAAGCAAGATG 441  
 QY 141 RepSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSer 160  
 Db 442 GATTCCAGCCACCAAGAAAGATCTTCAAGAGACCCCGCAGCAGCGGACCGTGAAGC 501  
 QY 161 ArgAspLeuCysHisMetAlaAsnGlyTleGluAspCysAspAspGlnProProLysGlu 180  
 Db 502 CGTGAATTTATGTACATGGCAATGGGATTTGAAGCTGTGATGATGATGATGATGATG 561  
 QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200  
 Db 562 AAGAAATCCAGTCAGCAAGAAAAAGAAACGCTCCAGGGCAAGCAGGAAAGGGAAGCT 621  
 QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220  
 Db 622 TCACCTGTGTAGTTTGCATAGACTCTAATGAACCTACATACCTGTTATGCAACCCAGTG 681  
 QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluThrPheHis-Ph 240  
 Db 682 TCTTATGGGAGATGATAGATGTCACATGA-CHGTGTCCAAATGGAATGGGTTCACCTT 740  
 QY 240 eSerCysValSerLeuThrTyrLys-ProLysGlyLysTyrTyr-CysProLysCysArg 259  
 Db 741 TTCATGTGTTCACCTTACCTATAAACCCCAAGGGGAAATGGTATTGTCACCAAGTCAGG 800  
 QY 260 -GlyAspAsn-----GluLysThrMetAspLysSerThrGluLysTh 273  
 Db 801 GGGAGATATGGAGAAAAAACCATGGGACAAAAAAGTACTGGAAAAAGGACCAAAAAAGGAA 860  
 QY 273 rLysLysAspArgArg 278  
 Db 861 TAGAAAAGATCCAAAG 876  
 RESULT 5  
 BQ225297  
 LOCUS 603799281F1 CSEQCHN23 Gallus gallus cdna clone ChEST768a21 5', mRNA  
 DEFINITION sequence.  
 ACCESSION BQ225297  
 VERSION BQ225297.1 GI:25461055  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 736)  
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Pong, W.T., Tickle, C., Brown, W.R.A.; Wilson, S.A. and Hubbard, S.J.  
 TITLE A Comprehensive Collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 22335534  
 PUBMED 12445392

## COMMENT

Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

## FEATURES

source

1. 736  
 Location/Qualifiers  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST768a21"  
 /dev\_stage="22"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQCHN23"  
 /note="Organ: heads; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996) 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 242 a 163 c 191 g 140 t

## ORIGIN

## Alignment Scores:

Pred. No.: 4,396-63 Length: 736  
 Score: 1123.00 Matches: 208  
 Percent Similarity: 92.98% Conservative: 17  
 Best Local Similarity: 85.95% Mismatches: 17  
 Query Match: 75.83% Indels: 0  
 DB: 13 Gaps: 0

US-09-513-365A-1 (1-280) x BU225297 (1-736)

QY 39 ProHisaspMetGlnArgasnValSerValLeuArgGluLeuaspAsnLysTyrGlnGlu 58  
 DB 3 CCGCTGAAGGAATACAGCGCTCTATGAAAAATACAGTCTGAGAACGATCCGCTCAG 62  
 QY 59 ThrLeuLysGluLeuaspValTyrGluLysTyrLysGluaspAspLeuaspGln 78  
 DB 63 GCTTTAAGGAATACAGCGCTCTATGAAAAATACAGTCTGAGAACGATCCGCTCAG 122  
 QY 79 LysLysArgGluGlnLeuLeuGlnArgAlaLeuLeuaspSerGlnGluLeuGlyasp 98  
 DB 123 AAGAAACGCTTGCCAGCAGCCTCCAGCGGCTTTAATCAACAGTCAAGAACTCGGAGAT 182  
 QY 99 GluLysIleGlnIleValThrGlnMetLeuGluLeuValGluasnArgAlaArgGlnMet 118  
 DB 183 GAGAAATTCAGATAGTACTCAGATCTGGAATCTGTAGAGAAATAGAGCCGACAGATG 242  
 QY 119 GluLeuHisSerGlnCysPheGlnaspProAlaGluSerGluArgAlaSeraspLysAla 138  
 DB 243 GAACACACTCTCAGTGTCTTTCAGGATCTGCTGAAAACGACAAAGCTCTGGAAGGCC 302  
 QY 139 LysMetaspSerSerGlnProGluArgSerArgProArgArgGlnArgThrSer 158  
 DB 303 AAGATGGAGTCTGCGCAGCCAGAGATCATCTACTAGACCTCGTGGCAGGAAACGAGC 362  
 QY 159 GluSerArgaspLeuGluCysHisMetAlaasnGlyIleGluaspCysaspGlnProPro 178  
 DB 363 GAGAGCGGACCTGTGGCATATAGCCACGCGGATGCTGACTGCGACGATCAGCCACCT 422

QY 179 LysGluLysLysSerLysSerLysLysLysLysLysLysLysLysLysLysLysLysLys 198  
 DB 423 AAAGAGAAAAGATCGAAATCTTCCAAAGAAAGAAAACGCTCCAAAGCCAAACAGAGAGG 482  
 QY 199 GluAlaSerProValGluPheAlaIleAspProasnGluProThrTyrCysLeuGlnCysasn 218  
 DB 483 GAGGTTTCCACCGTGGAGTTTGGATTTGATCCCAATGAACCCGACTTACTGCTTATGTAAC 542  
 QY 219 GlnValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluIlePhe 238  
 DB 543 CRAAGTGTCTTACGGCGAATGATAGATGTGATACGACAGTGTCTTATGAGTGGTTC 602  
 QY 239 HisPheSerCysValSerLeuThrTyrLysProLysGlyLysTyrCysProLysCys 258  
 DB 603 CACTTCTCGTGTGTGGACTCACCTATAAACCCGAGGGGAAATGGTATTGCCCAAGTGC 662  
 QY 259 ArgGlyAspAsnGluLysThrMetaspLysSerThrGluLysThrLysLysAspArgArg 278  
 DB 663 AGAGGAGATACGAGAAAACGATGGACAAATGTACTGACAAATCAAAAAAGGATAGAAGA 722  
 QY 279 SerArg 280  
 DB 723 TCGAGG 728  
 RESULT 6  
 BU332089  
 LOCUS  
 DEFINITION  
 sequence.  
 ACCESSION  
 BU332089  
 VERSION  
 BU332089.1 GI:25840090  
 KEYWORDS  
 EST.  
 SOURCE  
 Gallus gallus (chicken)  
 ORGANISM  
 Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 MEDLINE  
 PUBMED  
 12445392  
 CONTACT: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1. .892  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST94a23"  
 /tissue\_type="whole embryo"  
 /dev\_stage="10"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQCHN65"  
 /note="Organ: whole embryo; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2

## FEATURES

source

rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 284 a 190 c 224 g 194 t  
ORIGIN

## Alignment Scores:

Pred. No.: 6.8e-63 Length: 892  
Score: 1121.00 Matches: 210  
Percent Similarity: 93.06% Conservative: 18  
Best Local Similarity: 85.71% Mismatches: 16  
Query Match: 75.69% Indels: 1  
DB: 13 Gaps: 0

US-09-513-365A-1 (1-280) x BU332089 (1-892)

QY 37 SerLeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAenLysTyr 56  
DB 12 TCCTGCGCTGACATCCAGCGCAAGTGTCTGCTGCGGAGGTGGACACCCCGTGC 71  
QY 57 GlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspLeu 76  
DB 72 CAAGAAGCTTTAAAGCAATAGACGACCTCTATGAAAAATACAAAGTCTGAGAAGCATCT 131  
QY 77 AsnGlnLysLysArgGlnGlnLeuLeuGlnArgAlaLeuLeuLeuSerGlnGluLeu 96  
DB 132 GCTCAGAAGAAACGCTTGACAGACACCTCCAGCGGCTTTATCAACAGTCAAGAACATC 191  
QY 97 GlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArg 116  
DB 192 GGAGATGAGAAATTCAGATGATTATCAGTCTGGAAGTGTAGAGATAGAGCCCGA 251  
QY 117 GlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAsp 136  
DB 252 CAGATGGAAACACACTCTCAGTGTCTGAGTCTGCTGAAACGACAAAGCTCTGGAA 311  
QY 137 LysAlaLysMetAspSerSerGlnProGluArgSerSerArgArgProArgGlnArg 156  
DB 312 AAGCCCAAGATGAGGTCCTGCCAGCCAGAGAGATCATCAGTAGACCTCGTCGGCAGCGA 371  
QY 157 ThrSerGluSerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGln 176  
DB 372 ACCAGCGAGAGCGCGACCTGTCCATATAGCCAAACGGATCGATGACTGCGACGATCAG 431  
QY 177 ProProLysGluLysLysSerLysSerAlaLysLysLysArgSerLysAlaLysGln 196  
DB 432 CCACCTAAAGAGAAAGATCGAAATCTTCAAGAGAAACGCTCCAAAGCCAAACAA 491  
QY 197 GluArgGluAlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeu 216  
DB 492 GAGAGGGAGGTTTACCCTGGGAGTTTGGCATTTGATCCCAATGAACCGACTTACTGCTTA 551  
QY 217 CysAsnGlnValSerTyrGlyGluMetIleGlyCysAsp-AsnGluGlnCysProIleG1 236  
DB 552 TGTAACCAAGTGTCTACGCGGAATGATAGGATGTGATTAACGACAGTGTCTATTGA 611  
QY 236 uTPpHeHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTrpTyrCysPr 256  
DB 612 GTGTTTCCACTTCTCGTGTGTGGACTCACTATAAACCGAGGGGAAATGGTATTGCC 671  
QY 256 oLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAs 276  
DB 672 CAAGTCGAGAGAGATAACGAGAAACGATGGCAAAATGTACTGACAAATCAAAACAGGA 731  
QY 276 pArgArgSerArg 280  
DB 732 TAGAAGATCGAGG 744

## RESULT 7

LOCUS BG211544

DEFINITION RST31111 Atherysys RAGE Library Homo sapiens cDNA, mRNA sequence.  
EST 21-APR-2001

ACCESSION BG211544  
VERSION BG211544.1 GI:13733231  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

## REFERENCE

1 (bases 1 to 919)  
Harrington,J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,  
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,  
Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,  
E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,  
J., Danzig, J., and Ducar, M.

## TITLE

Creation of genome-wide protein expression libraries using random  
activation of gene expression

JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
PUBMED 11329013

## COMMENT

Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave., Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@athersys.com  
High quality sequence stop: 439.

## FEATURES

source

1..919  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression'  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

BASE COUNT 332 a 151 c 216 g 216 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.6e-61 Length: 919  
Score: 1096.50 Matches: 214  
Percent Similarity: 96.00% Conservative: 2  
Best Local Similarity: 95.11% Mismatches: 7  
Query Match: 74.04% Indels: 4  
DB: 10 Gaps: 1

US-09-513-365A-1 (1-280) x BG211544 (1-919)

QY 58 GluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspLeuAsn 77  
DB 4 GAAAGCTTAAAGAAATTCATGATGCTACGAAAAATATAAGAAAGATGATGTTTAAAC 63  
QY 78 GlnLysLysArgLeuGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGly 97  
DB 64 CAGAAGAAACGCTACAGCAGCTTCTCCAGAGAGACACTAATTAATAGTCAAGATTGGGA 123  
QY 98 AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln 117  
DB 124 GATGAAAAAATACAGATTGTTACACAAATGCTCGAATTCGTGGAAATCGGCAACACAA 183  
QY 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLys 137  
DB 184 ATGGAGTTACACTCAGAGTGTTCGAAGATCTCTGCTGAAAGTCAACGAGCCTCAGATAA 243  
QY 138 AlaLysMetAspSerSerGlnProGluArgSerArgArgProArgGlnArgThr 157  
DB 244 GCAAGATGATGATTCAGCCCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGGCGGACC 303  
QY 158 SerGluSerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspGlnPro 177





JOURNAL  
COMMENT

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11661 row: n column: 13  
High quality sequence stop: 775.

## FEATURES

Location/Qualifiers  
1..778

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5262540"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/clone\_lib="NIH MGC\_95"

/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTCTTTTCTTTTCTT-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 237 a 176 c 227 g 138 t

## ORIGIN

## Alignment Scores:

Pred. No.: 3,25e-59 Length: 778  
Score: 1063.00 Matches: 215  
Percent Similarity: 98.63% Conservative: 1  
Best Local Similarity: 98.17% Mismatches: 3  
Query Match: 71.78% Indels: 2  
DB: 12 Gaps: 0

US-09-513-365A-1 (1-280) x BI548536 (1-778)

QY 1 MetLeuGlyClnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20  
DB 124 ATGTTAGGCGAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTGACCGGGAGCG 183  
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40  
DB 184 AGCGGCTGCTCACCTGCTACGTGCGAGGACTACCTTGAGTGGTGCGTGCACAC 243  
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60  
DB 244 GACATGCAGAGGAACGTGCTGCTGCGAGAGCTGGACACAAATATCAGAAAGTTA 303  
QY 61 LysGluIleAspAspValTyrGluLysTyrLysGluAspAspLeuAsnGlnLysLys 80  
DB 304 AAGGAATTCATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCAAGAGAAA 363  
QY 81 ArgLeuGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys 100  
DB 364 CGTCTACAGAGCTTCTCAGAGAGCACTAATTAATAGTCAAGAAATGGAGATGAAAA 423  
QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120  
DB 424 ATACAGATTGTACACAAATGCTCGAATGTTGTTGAAATCGGACAGACAAATGGAGTTA 483  
QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140  
DB 484 CACTCACAGTGTTCAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAGCAAGATG 543  
QY 141 AspSerSerGlnProGluArgSerSerArgArgProArgGlnArgThrSerGluSer 160

DB 544 GATTCACGCCAACCAAGATCTTTCAAGAACACCCCGCAGCAGCGGACCAAGTGAAGC 603  
QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180  
DB 604 COTGATTATGTACATGCAATGGATTGAAGACTGTGATGATGATGATGATGATGATGAT 663  
QY 181 LysLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLys 200  
DB 664 AAGAATTCAGTTCAGCAAGAAACGCTCAAGGC-AAGCAGGAAGAGGAGT 722  
QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGln 219  
DB 723 TCACCTGTTGAGTTTGC-ATAGATCTTAATGAACCTACATACCTGCTTATGCAACAAG 778

## RESULT 10

BU259874 779 bp mRNA linear EST 26-NOV-2002  
LOCUS 603504254F1 CSEQCHN51 Gallus gallus cDNA clone CHEST426k14 5', mRNA  
DEFINITION sequence.

ACCESSION BU259874 GI:25526963

VERSION BU259874

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 779)

REFERENCE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

JOURNAL MEDLINE

PUBMED 12445392

COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..779

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="White Leghorn, Hise"

/db\_xref="taxon:9031"

/clone="CHEST426k14"

/dev\_stage="22"

/lab\_host="DH10B"

/clone\_lib="CSEQCHN51"

/note="Organ: limbs; Vector: pBluescript II KS(+); Site\_1:

ECORI; Site\_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunted, ligated to NotI adapters, digested with EcoRI

size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

BASE COUNT 261 a 155 c 191 g 171 t

ORIGIN

Alignment Scores:

Pred. No.: 6,25e-58 Length: 779

Score: 1043.00 Matches: 193

```
Percent Similarity: 94.06% Conservative: 13
Best Local Similarity: 88.13% Mismatches: 13
Query Match: 70.43% Indels: 0
DB: 13 Gaps: 0

US-09-513-365A-1 (1-280) x BU259874 (1-779)

QY 62 GluileAspValTyrGluLysTyrLysLysGluAspLeuAenGlnLysLysArg 81
Db 2 GAAATAGACGACGCTATGAAAAATACAACTGTGAGAACGATCTGCTCAGAGAAACGC 61
QY 82 LeuGlnGlnLeuGlnArgAlaLeuLeuAenSerGlnGlnLeuGlnLysLysLys 101
Db 62 TTGACGACGACCTCCAGCGGGCTTTAATCAACAGTCAAGAACTCGGAGATGAGAAAT 121
QY 102 GlnileValThrGlnMetLeuGlnLeuValGluAenArgAlaArgGlnMetGlnLeuHis 121
Db 122 CAGATAGTACTCAGATGCTGGAACTGTGAGAAATAGACCCGACAGATGGAACACAC 181
QY 122 SerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMetAsp 141
Db 182 TCTCAGTGTCTTTCAGGATCTGTGAAACGACAGCGCTCTGGAAGGCGCAAGATGGAG 241
QY 142 SerSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSerArg 161
Db 242 TCTGCGCAGCCAGAGATCATCAGCTAGACCTCGTGGCAGCGAAACAGCGAGAGCCGC 301
QY 162 AspLeuCysHisMetAlaAenGlyLeuGluAspCysAspAenGlnProProLysGlnLys 181
Db 302 GACCTGTGCATATAGCAACGGGATCGATGATCGCAGCATCAGCCACCTAAAGAGAA 361
QY 182 LysSerLysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSer 201
Db 362 AGATCGAATCTTCCAGAGAAACAAACGCTCCAAAGCCAAACAGAGAGGGAGTTTCA 421
QY 202 ProValGlnPheAlaLeuAspProAenGluProThrTyrCysLeuCysAenGlnValSer 221
Db 422 CCGGTGGAGTTTGGATGATGCTCAATGAAACCGACTTACTGCTTATGTAACCAAGTGT 481
QY 222 TyrGlyClnMetileGlyCysAenGlnGlnCysProileGlnThrPheHisPheSer 241
Db 482 TAGCGGAAATGATAGATGATGATAACGAAACAGTGTCTTATGAGTGGTTTCCACTTCT 541
QY 242 CysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAsp 261
Db 542 TGTGTGGAGTCACTATATAACGAGGGAATGATTTGGCCCAAGTGCAGAGAGAT 601
QY 262 AsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg 280
Db 602 AACGAGAAACGATGACAAATGATCTGACAAATCAAAAAGGATAGAGATCGAGG 658

RESULT 11
BU386788 673 bp mRNA linear EST 28-NOV-2002
LOCUS 603581891F1 CSQCHN75 Gallus gallus cdna clone CHEST531f23 5', mRNA
DEFINITION sequence.
ACCESSION BU386788
VERSION BU386788.1 GI:25894789
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 673)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
```

```
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk;
Location/Qualifiers
1. .673
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST531f23"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSQCHN75"
/note="Organ: trunks; Vector: pBluescript II KS(+);
Site_1: EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
BASE COUNT 206 a 157 c 179 g 131 t
ORIGIN
Alignment Scores:
Pred. No.: 5,978-57 Length: 673
Score: 1027.00 Matches: 193
Percent Similarity: 92.86% Conservative: 15
Best Local Similarity: 86.16% Mismatches: 15
Query Match: 69.35% Indels: 1
DB: 13 Gaps: 0
US-09-513-365A-1 (1-280) x BU386788 (1-673)
QY 33 GluLysValGluSerLeuProHis-AspMetGlnArgAenValSerValLeuArgGluLe 52
Db 2 GAGTGGTGGAGTCTGCTGGCGCTGGACATCCAGGCGCAACGTCGCTGCTCGGGAGGT 61
QY 52 uAspAenLysTyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLys 72
Db 62 GCACACCCCGTCCCAAGAGCTTTAAAGGAAATAGACGACGCTCTATGAAAAATACAAGTC 121
QY 72 sGluAspLeuAenGlnLysLysArgLeuGlnGlnLeuGlnArgAlaLeuLeuAs 92
Db 122 TGAGAACGATCTGCTGCAAGAGAAACGCTTGCAGCAGCAGCACCTCCAGCGGGCTTATCAA 181
QY 92 nSerGlnGlnLeuGlyAspGluLysLeuGlnLeuValThrGlnMetLeuGlnValG 112
Db 182 CAGTCAAGAACTCGGAGATGAGAAATTCAGATAGTTACTCAGATCTCGAATCTGGTAGA 241
QY 112 uAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerG 132
Db 242 GAATAGAGCCCGACAGATGGAACACACTCTCAGTGTCTTTCAGGATCTCTCTGAAACCGA 301
QY 132 uArgAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerSerArgArgPr 152
Db 302 CAAGCCTCTGGAAGAGGCGCAAGATGGAGTCTCTGCCAGCCAGAGAGATCATCAGTAGACC 361
QY 152 oArgArgGlnArgThrSerGluSerArgAspLeuCysHisMetAlaAenGlyLeuLys 172
Db 362 TCGTGGCAGCGAAACAGCGAGCGCGACCTGTGCCATATAGCAACGGAGATCGATGA 421
QY 172 pCysAspLeuGlnProProLysLysSerLysSerAlaLysLysLysLysArgSe 192
```

Db 422 CTGCGACGATCAGCCACTTAAGAGAGAAAGATCGAAATCTTCCAGAGAAAAACGCTC 481  
 Qy 192 rLysAlaLysGlnAlaArgGluAlaSerProValGluPheAlaIleAspProAsnGluPr 212  
 Db 482 CAAAGCCAAACAAGAGAGGAGGTTTACCCTGGAGTTTGCCATTGATCCCAATGAACC 541  
 Qy 212 oThrTyCysLeuCysAsnGlnValSerTyGlyGluMetIleGlyCysAspAsnGluGl 232  
 Db 542 GACTTACTGTTATGTAACCAAGTCTTACGGGAAATGATAGGATGTGATTAACGACA 601  
 Qy 232 nCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyLysProLysGlyLy 252  
 Db 602 GTGCTCTATTAGTGGTTCACCTTCGTGTGTGGACTCACCTATAAACCGAAGGGAA 661  
 Qy 252 sTrpTyCys 255  
 Db 662 ATGGTATTGC 671

## RESULT 12

BM982877/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BM982877 735 bp mRNA linear EST 21-FEB-2003  
 UI-CF-EN1-acs-d-05-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone  
 UI-CF-EN1-acs-d-05-0-UI 3', mRNA sequence.

BM982877  
 BM982877.1 GI:19606826  
 EST.  
 Homo sapiens (human)

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS  
 TITLE

Bonaldo M.F., Lennon G. and Soares M.B.  
 Normalization and subtraction: two approaches to facilitate gene

discovery  
 Genome Res. 6 (9), 791-806 (1996)

JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT

97044477  
 889548  
 Contact: McCray, PB

McCrack Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866  
 Fax: 319 356 7171

Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems  
 (www.openbiosystems.com).

Seq primer: M13 FORWARD  
 POLYA=Yes.

## FEATURES

source  
 Location/Qualifiers  
 1..735  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EN1-acs-d-05-0-UI"  
 /tissue\_type="Primary Lung Cystic Fibrosis Epithelial  
 Cells"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-EN1"  
 /notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-CF-EN1 is a normalized cDNA library containing the  
 following tissue(s): Primary Lung Cystic Fibrosis  
 Epithelial Cells. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not

I, and cloned directionally into pT7T3-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is CTGCTCAGGT.

TAG LiB=UI-CF-EN1

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS

6hr To LPS 24h

TAG\_SEQ=CTGCTCAGGT"

BASE COUNT 180 a 161 c 114 g 280 t

ORIGIN

Alignment Scores:

Pred. No.: 9.93e-57 Length: 735

Score: 1024.00 Matches: 190

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 69.14% Indels: 0

DB: 12 Gaps: 0

US-09-513-365A-1 (1-280) x BM982877 (1-735)

Qy 91 IleAsnSerGlnGluLeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeu 110

Db 735 ATTAATAGTCAAGAAATGGAGATGAAAAATACAGATTGTTACACAAATGCTCGAATTG 676

Qy 111 ValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu 130

Db 675 GTGGAAAAATCGGGCAAGACAAATGAGTTACACTCACAGTGTTCACAGATCTCTGCTGAA 616

Qy 131 SerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerSerArg 150

Db 615 AGTGACGAGCTTCAGATAAAGCAAGATGGATTCCAGCCACCAAGAGATCTTCAAGA 556

Qy 151 ArgProArgArgGlnArgThrSerGluSerArgAspLeuGluHisMetAlaSerGlyIle 170

Db 555 AGACCCCGCAGCGAGCGAGCCAGTGAAGCCGTGATTTATGTACATGGCAATGGGATT 496

Qy 171 GluAspCysAspAspGlnProLysGluLysSerLysSerLysSerLysLysLysLys 190

Db 495 GAAGACTGTGATGATCAGCCACTTAAGAAAGAAATCAAGTCAGCAAGAAAGAAAGAAA 436

Qy 191 ArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIleAspProAsn 210

Db 435 CCTTCAAGCCCAAGCAGAAAGGAGCTTCACTGTGAGTTGCAATAGATCTCTAAT 376

Qy 211 GluProThrTyCysLeuCysAsnGlnValSerTyGlyGluMetIleGlyCysAspAsn 230

Db 375 GAACCTACATACCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTGACAAT 316

Qy 231 GluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyLysProLys 250

Db 315 GAACAGTGTCCAATTAATGGTTTTCATCTTTCATGTTTTCATCTTACCTATAAACCAAG 256

Qy 251 GlyLysTrpTyCysProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThr 270

Db 255 GGGAAATGTTATGCCCCAAGGAGGAGGAGATATAGAGAAACAATGGCAACAAAGTACT 196

Qy 271 GluLysThrLysLysAspArgArgSerArg 280

Db 195 GAAAGACAAAAAGGATAGAGATCGAGG 166

RESULT 13

BU323498

LOCUS

DEFINITION

BU323498 707 bp mRNA linear EST 28-NOV-2002

603489003F1 CSEQCHN63 Gallus gallus cDNA clone ChEST388e2 5', mRNA

sequence.

ACCESSION

BU323498

VERSION

BU323498.1 GI:25831499

KEYWORDS

EST.

SOURCE

Gallus gallus (chicken)

ORGANISM

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.

## REFERENCE

## AUTHORS

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
A Comprehensive Collection of Chicken cDNAs  
Curr. Biol. 12 (22), 1965-1969 (2002)

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

12445392

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

## FEATURES

source

Location/Qualifiers

1..707

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="White Leghorn, Hisex"

/db\_xref="taxon:9031"

/clone="CHEST388e2"

/dev\_stage="36"

/lab\_host="DH10B"

/clone\_lib="CSECHN63"

/note="Organ: heads; Vector: pBluescript II KS(+); Site\_1:  
EcORI; Site\_2: NotI; This normalized library was  
constructed from 1 million independent clones. cDNA  
synthesis was initiated using an oligo(dT) primer, using  
methylation C in the first strand synthesis reaction.  
Following this first strand reaction, double-stranded cDNA  
was bluntly ligated to NotI adapters, digested with EcORI  
, size-selected, and cloned into the NotI and EcORI  
compatible sites of a custom modified MCS of the  
pBluescript (KS+) vector. The library was normalized in 2  
rounds using conditions adapted from Soares et al., PNAS  
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
(1996): 791, except that a significantly longer  
reannealing hybridization was used."

205 a 171 c 195 g 136 t

## BASE COUNT

## ORIGIN

Alignment Scores:  
Pred. No.: 1.3e-56 Length: 707  
Score: 1022.00 Matches: 195  
Percent Similarity: 91.77% Conservative: 17  
Best Local Similarity: 84.42% Mismatches: 18  
Query Match: 69.01% Indels: 1  
DB: 13 Gaps: 0

US-09-513-365A-1 (1-280) x BUS23498 (1-707)

QY 18 GYGLuArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37  
Db 15 GGGGAGCGGCGCGGCTGCTTCCTAGCTAGTCCAGGACTACTGAGTGGTGGATCG 74  
QY 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln 57  
Db 75 CTGCGGCTGACATCCAGCGCAACGTCCTGCTCGCGGAGTGGACACCGCGTGCCAA 134  
QY 58 GluThrLeuLys-GluLeuAspValTyrGluLysTyrLysLysGluAspAspLeuAs 77  
Db 135 GAAGCTTTAAAGGGAATAGACACGCTCTATGAAATAATACAGTCTGAGAACGATCCTGC 194  
QY 77 nGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuG 97  
Db 195 TCAGAGAACCGTTTCAGAGAGACCTCCAGCGGGCTTTATCAACAGTCAAGACTTGG 254  
QY 97 YAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgG 117  
Db 255 AGATGAGAAATTCAGATAGTACTTACTCAGATGCTGGAACCTGGTAGAATAGAGCCCGACA 314

QY 117 nMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLy 137  
Db 315 GATGAAACACACTCTCAGTGTCTGAGGATCTGCTGAAACACGCAAGCCTCTGCAAAA 374  
QY 137 sAlaLysMetAspSerSerGlnProGluArgSerArgArgProArgArgGlnArgTh 157  
Db 375 GGCCAAGATGGAGTCTCTCCAGCCAGAGAGATCATCAGTAGACTCTGTCGCGAGCAAC 434  
QY 157 rSerGluSerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnPr 177  
Db 435 CAGCAGAGAGCCCGACCTGTGCTATATAGCCAGCGATCGATGATCGACATCAGCC 494  
QY 177 oProLysGluLysLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLys 197  
Db 495 ACCTAAAGAGAAAGATCGAAATCTTCCAAGAAGAAAGAAAGCTCCAAAGCAACCAAGA 554  
QY 197 uArgGluAlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCy 217  
Db 555 GAGGAGAGTTTCACCCGTGGAGCTTGGATTCATCCATGAACCGACTTACTGCTTATG 614  
QY 217 sAsnGlnValSerTyrGluMetIleGlyCysAspAsnGluGlnCysProIleGluTr 237  
Db 615 TAAACAAGTGTCTTACGGCGAAATGATAGGATGTATAACGAACAGTGTCTTATTGAGTG 674  
QY 237 pPheHisPheSerCysValSerLeuThrTyr 247  
Db 675 GTTCCACTTCTGCTGTAGGACTCACCTAT 705  
BUS101073 864 bp mRNA linear EST 20-SBP-2002  
AGENCOURT 8866651 NIH\_MGC\_137 Mus musculus cDNA clone IMAGE:6430926  
5', mRNA sequence.  
BUS101073  
BUS101073.1 GI:23242616  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gerard Gradwohl (PNAS 97 P1607-1611, 2000)  
cDNA Library Preparation: Catherine Lee, Endocrine Pancreas  
Consortium  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: IRBD11 row: a column: 07  
High quality sequence stop: 560.  
Location/Qualifiers  
1..864  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6430926"  
/lab\_host="DH10B"  
/clone\_lib="NIH MGC\_137"  
/note="Organ: pancreas; Vector: pSPOR1; Site 1: SalI;  
Site 2: NotI; Library consists of a pool of clones  
rearrayed from the following libraries: Melton normalized  
mixed mouse pancreas 1 NI-MMS1, Amplified Melton mouse  
islets 1 Misi-A, and Kaestner ngm3 wt. Clones rearrayed in  
the laboratory of K. Kaestner (University of Pennsylvania  
). Note: this is a NIH MGC Library."  
197 a 217 c 161 g 289 t  
BASE COUNT

## ORIGIN

Alignment Scores: 1.28e-55 Length: 864  
 Pred. No.: 1007.50 Matches: 197  
 Score: 1007.50  
 Percent Similarity: 91.23% Conservative: 11  
 Best Local Similarity: 86.40% Mismatches: 16  
 Query Match: 68.03% Indels: 5  
 DB: 13 Gaps: 2

US-09-513-365A-1 (1-280) x BU591073 (1-864)

QY 57 GlnGluThrLeuLysGlu-----lleAspAspValTyrGluLysTyrLys-LysGluAs 74  
 Db 841 CAGGAACCTTTAAAGGGAATTGAGTGCTTTGGAAAAATTTTAGGGAAGAGGG 782  
 QY 74 pAspLeuAan---GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSe 93  
 Db 781 TGGTTTCCAAACCAAGAAAAACCCCTCCAGCAGCATTTCCAGAGAGCGTTAATCAATAG 722  
 QY 93 xGlnGluLeuGluYAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAs 113  
 Db 721 CCAAGAAATTTGGGAGATGAAAAAATTCAGATTGCCCCAGATGCTGGAATTTGG-GAGAA 663  
 QY 113 nArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluAr 133  
 Db 662 CCGAGCGAGACAATCGAGCTGCATTCACAGTGTTCAGAGATCCTGCTGAAAGTGAGCG 603  
 QY 133 gAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerSerArgProAr 153  
 Db 602 AGCCTCAGACAAGTCGAAGATGATTCAGTCAACCGGAAGATCTTCTAGAAGACCTCG 543  
 QY 153 gArgGlnArgThrSerGlnSerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCy 173  
 Db 542 AAGACAGAGACAGTGAGAGCGGTGACTATGTGCATGCACAAACGGGATTCAGACTG 483  
 QY 173 sAspAspGlnProProLysGluLysLysSerLysSerAlaLysLysLysLysArgSerLy 193  
 Db 482 TGNATGATCAACCCCGAAGAAAGAGATCCAAAGTCCGCAAGAAAGAGCGCTCCAA 423  
 QY 193 sAlaLysGlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGluProth 213  
 Db 422 GGCACAGCAGGAGAGGAGCATCCCTGTGCGAGTTTGCCATGCATCCCAATGAGCCAC 363  
 QY 213 rTyrCysLysCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCy 233  
 Db 362 CTACTGCTTGTGTAAACCAAGTGTCTACGGGGAGATGATAGGCTGTGACAATGAACAGTG 303  
 QY 233 sProfileGluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTr 253  
 Db 302 TCCCATTTGAATGGTTTCATCTTTTCATGTGTTTCATCTTAAACCCNAGGGGAATG 243  
 QY 253 pTyrCysProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGluLysTh 273  
 Db 242 GTATTCCCAAAAGTGTAGGGAGACAATGAGAAAAACCATGGACAAAAGTACCGAAAAGAC 183  
 QY 273 rLysLysAspArgArgSerArg 280  
 Db 182 AAAAAAGGAGAAGAGCGAGG 161

## RESULT 15

BG184056/c

LOCUS BG184056 798 bp mRNA linear EST 21-APR-2001  
 DEFINITION RST2972 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG184056  
 VERSION BG184056.1 GI:13705743  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 798)  
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,

Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,  
 Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith  
 ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher  
 ,J., Danzig,J. and Ducar,M.  
 Creation of genome-wide protein expression libraries using random  
 activation of gene expression

Nat. Biotechnol. 19 (5), 440-445 (2001)  
 2127151  
 PUBMED

COMMENT  
 11329013  
 Contact: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@atersys.com

High quality sequence stop: 531.  
 Location/Qualifiers  
 1..798  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="HT1080"  
 /clone\_lib="Athersys RAGE Library"  
 /note="See 'Creation of Genome-wide Protein Expression'  
 Libraries using Random Activation of Gene Expression',  
 Nature Biotechnology, in press. Note that even though the  
 cell type indicated is HT1080, since a random activation  
 method was used, these sequence tags are not necessarily  
 expressed in HT1080 under normal circumstances."

FEATURES  
 source

BASE COUNT 197 a 178 c 130 g 293 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.4e-55 Length: 798  
 Score: 1006.50 Matches: 209  
 Percent Similarity: 95.43% Conservative: 0  
 Best Local Similarity: 95.43% Mismatches: 10  
 Query Match: 67.96% Indels: 6  
 DB: 10 Gaps: 0

US-09-513-365A-1 (1-280) x BG184056 (1-798)

QY 62 GluIleAspAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLysArg 81  
 Db 798 GAAATTTGATAT-GTCTACGCAAAATATAG-AAAGAAAGATGATTTAACCCAGAAAGACGT 741  
 QY 82 LeuGlnGlnLeuLeuGlnArgAlaLeuLeuLeuSerGlnGlnLeuGluGlyAspGluLysIle 101  
 Db 740 CTACAGCAGCTTCTCCAGAGAGCACTATTAT-AGTCAAGATGGGAGATGAAAAA- 683  
 QY 102 GlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeuHis 121  
 Db 682 CAGATTGTTC-CAAAATGCTCGAATTTGGTGA-ANTCGGGCAAGACAAATGAGTTTACAC 625  
 QY 122 SerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMetAsp 141  
 Db 624 TCACAGTGTTCACAGATCCTGCTGAAAGTGAACGAGCTCAGATTAACCAAGATGGAT. 565  
 QY 142 SerSerGlnProGluArgSerSerArgProArgGlnArgThrSerGluSerArg 161  
 Db 564 TCCAGCCCAACACAGAAAGATCTTCAAGAGAGACCCCGCAGCGAGACAGTGAAGCCGT 505  
 QY 162 AspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLys 181  
 Db 504 GATTTATGTGCATGCGCAATGGGATTGAAGACTCTGTGATGATCAGCCACCTAAAGAAAAG 445  
 QY 182 LysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGlnGluArgGluAlaSer 201  
 Db 444 AAATCCAGTTCAGCAAGAAAGAAACGCTCCAGGGCCAGCAGAAAGGAGGAGTTCA 385  
 QY 202 ProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSer 221  
 Db 384 CTTGTTGAGTTTGCAATAGATCTCTAATGAACCTACATACCTACTGCTTATGCAACCAAGTGTCT.325

```
Qy 222 TyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSer 241
    |||||
Db 324 TATGGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTGAATGGTTTCACTTTCA 265
    |||||
Qy 242 CysValSerLeuThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAsp 261
    |||||
Db 264 TGTGTTTTCACCTTACCTATAAACCAAGGGGAAATGGTATTGCCCAAAAGTCAGGGGAGAT 205
    |||||
Qy 262 AsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg 280
    |||||
Db 204 AATCAGAAAACAATGGACAAAAGTACTGAAAAGACAAAAAGGATAGAGATCGAGG 148
    |||||
```

Search completed: January 12, 2004, 20:28:52  
Job time : 2038 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2004, 19:24:44 ; Search time 77 Seconds  
(without alignments)  
1605.030 Million cell updates/sec

Title: US-09-513-365A-1  
Perfect score: 1481  
Sequence: 1 MUGQQQQQLYSSAALLTGR.....DNEKTMDSKTEKTKDRSR 280

Scoring table: BLOSUM62 Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09513365/runat\_12012004\_163932\_28784/app\_query.fasta\_1.455  
-DB=issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513365 @CNC 1.1 103 @runat\_12012004\_163932\_28784 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/2/ina/SA\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/SA\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/SA\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1481	100.0	840	US-09-601-478-6	Sequence 6, Appli
2	1481	100.0	1078	US-09-601-478-7	Sequence 7, Appli
3	1334	90.1	1154	US-09-484-970B-81	Sequence 81, Appli
4	853.5	57.6	873	US-09-006-783A-4	Sequence 4, Appli
5	728	49.2	1902	US-09-258-257-1	Sequence 1, Appli
6	728	49.2	1902	US-09-258-371-1	Sequence 1, Appli
7	728	49.2	1902	US-08-569-721A-1	Sequence 1, Appli
8	728	49.2	1902	US-08-751-230-1	Sequence 1, Appli
9	728	49.2	1902	US-09-499-082-1	Sequence 1, Appli
10	728	49.2	1902	US-09-258-372-1	Sequence 1, Appli
11	728	49.2	2061	US-09-258-371-9	Sequence 9, Appli
12	728	49.2	2061	US-08-751-230-9	Sequence 9, Appli

13	728	49.2	2061	3	US-09-499-082-9	Sequence 9, Appli
14	728	49.2	2061	3	US-09-258-372-9	Sequence 9, Appli
15	728	49.2	2061	4	US-09-159-871-1	Sequence 1, Appli
16	722	48.8	2061	3	US-09-006-783A-2	Sequence 2, Appli
17	662	44.7	633	3	US-09-006-783A-6	Sequence 6, Appli
18	400	27.0	699	3	US-09-195-286-2	Sequence 2, Appli
19	396.5	26.8	678	3	US-09-195-286-3	Sequence 3, Appli
20	359	24.2	1864	4	US-09-620-312D-435	Sequence 435, App
21	314	21.2	451	4	US-09-370-838-146	Sequence 146, App
22	172.5	11.6	8931	3	US-09-051-019-1	Sequence 1, Appli
23	127.5	8.6	1820	3	US-08-845-258-17	Sequence 17, Appli
24	127.5	8.6	1820	3	US-08-845-258-37	Sequence 37, Appli
25	127.5	8.6	1820	3	US-08-990-571-17	Sequence 17, Appli
26	127.5	8.6	1820	3	US-08-990-571-37	Sequence 37, Appli
27	127.5	8.6	1820	4	US-08-723-142A-17	Sequence 17, Appli
28	127.5	8.6	1820	4	US-08-723-142A-37	Sequence 37, Appli
29	127.5	8.6	1820	4	US-09-528-784A-17	Sequence 17, Appli
30	127.5	8.6	1820	4	US-09-528-784A-37	Sequence 37, Appli
31	127.5	8.6	1820	4	US-09-569-098A-17	Sequence 17, Appli
32	127.5	8.6	1820	4	US-09-569-098A-37	Sequence 37, Appli
33	127.5	8.6	2001	4	US-09-528-784A-84	Sequence 84, Appli
34	127.5	8.6	2001	4	US-09-569-098A-84	Sequence 84, Appli
35	127.5	8.6	3402	4	US-09-528-784A-86	Sequence 86, Appli
36	127.5	8.6	3402	4	US-09-569-098A-86	Sequence 86, Appli
37	125	8.4	6328	4	US-08-913-832A-1	Sequence 1, Appli
38	125	8.4	6328	4	US-09-249-181A-1	Sequence 1, Appli
39	125	8.4	6328	4	US-09-158-707-1	Sequence 1, Appli
40	125	8.4	6475	4	US-09-620-312D-325	Sequence 325, App
41	122	8.2	8600	4	US-09-457-708-1	Sequence 1, Appli
42	122	8.2	8600	4	US-09-950-046A-1	Sequence 1, Appli
43	120	8.1	1772	2	US-08-960-022-13	Sequence 13, Appli
44	118.5	8.0	2133	4	US-09-107-532A-1413	Sequence 1413, Ap
45	117.5	7.9	6314	4	US-09-620-312D-98	Sequence 98, Appli

ALIGNMENTS

RESULT 1  
US-09-601-478-6  
; Sequence 6, Application US/09601478  
; Patent No. 6403785  
; GENERAL INFORMATION:  
; APPLICANT: Ostuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human INGIL gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/09/601,478  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
US-09-601-478-6

Alignment Scores:  
Pred. No.: 6.8e-156 Length: 840  
Score: 1481.00 Matches: 280  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-513-365A-1 (1-280) x US-09-601-478-6 (1-840)

```
QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
Db 1 ATGTTAGGCGCAGCAGCAGCACTGTAATCGTCGGCGCGCTCTCGACCGGGAGCGG 60
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40
Db 61 AGCCGGCTGCTACCTGCTACGTCGAGGAGTACCTTGAAGTGGTGGAGTGGTGGCCAC 120
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
Db 121 GACATGACAGGAACGTCGTGCTGCGAGAGCTGGAACAACAATATCAAGAAACGTTA 180
QY 61 LysGluLeuAspAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLys 80
Db 181 AAGAAATTTGATGATGCTTACGAAATATATAGAAAGAGATGATTTAAACCAAGAGAA 240
QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlnLeuGlyAspGluLys 100
Db 241 CGTCTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATTTGGAGATGAAA 300
QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db 301 ATACAGATTTGTTACACAAATGCTCGAAATTTGGTGAATAATCGGCAAGACAAATGGAGTTA 360
QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
Db 361 CACTCACAGTGTTCACAGATCTCTGCTGAAGTGAACGAGCCTCAGATTAAGCAAGATG 420
QY 141 AspSerSerGlnProGlnArgSerSerArgProArgArgGlnArgThrSerGluSer 160
Db 421 GATTTCCAGCCCAACCAAGAGATCTTCAAGAGAGACCCCGCAGCGAGCCAGTGAAGC 480
QY 161 ArgAspLeuCysHisMetAlaLeuGlyLysGluAspCysAspAspGlnProProLysGlu 180
Db 481 CGTGATTTATGTACATGTCAAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAA 540
QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
Db 541 AAGAAATCCAGTTCAGCAAGAAAGAAACGCTCCAGCCCAAGCAGGAAGGAGGAGCT 600
QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db 601 TCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATCTGTTATGCAACCAAGTG 660
QY 221 SerTyrGlyClnMetIleGlyCysAspAsnGlnGlnCysProIleGluThrPheHisPhe 240
Db 661 TCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCCAATTTGAATGGTTTCACTTT 720
QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260
Db 721 TCATGTGTTTCACTTACCTTATACCAAGGAGGAAATGGTATTTGCCCAAGTGCAGGGA 780
QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg 280
Db 781 GATAATGAGAAACAAATGCAAAAGTACTGAAAGACAAAGAGATAGATGAGATCGAG 840
```

## RESULT 2

```
US-09-601-478-7
; Sequence 7, Application US/09601478
; Patent No. 6403785
; GENERAL INFORMATION:
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Human TSC403 gene and human INGIL gene
; FILE REFERENCE: Q60193
; CURRENT APPLICATION NUMBER: US/09601478
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP H10-134679
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: JP H10-73234
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP H10-38133
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 11
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human embryonic brain cDNA library
; NAME/KEY: CDS
; LOCATION: (92)..(931)
US-09-601-478-7
Alignment Scores:
Pred. No.: 9,83e-156 Length: 1078
Score: 1481.00 Matches: 280
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
```

```
US-09-513-365A-1 (1-280) x US-09-601-478-7 (1-1078)
QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
Db 92 ATGTTAGGCGCAGCAGCAGCACTGTAATCGTCGGCGCGCTCTCGACCGGGAGCGG 151
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40
Db 152 AGCCGGCTGCTACCTGCTACGTCGAGGAGTACCTTGAAGTGGTGGAGTGGTGGCCAC 211
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
Db 212 GACATGACAGGAACGTCGTGCTGCGAGAGCTGGAACAACAATATCAAGAAACGTTA 271
QY 61 LysGluLeuAspAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLys 80
Db 272 AAGAAATTTGATGATGCTTACGAAATATATAGAAAGAGATGATTTAAACCAAGAGAA 331
QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlnLeuGlyAspGluLys 100
Db 332 CGTCTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATTTGGAGATGAAA 391
QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db 392 ATACAGATTTGTTACACAAATGCTCGAATTTGGTGGAAATCGGCGAGACAAATGGAGTTA 451
QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
Db 452 CACTCACAGTGTTCCAAGATCTCTGCAAAAGTGAACGAGCCTCAGATTAAGCAAGATG 511
QY 141 AspSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSer 160
Db 512 GATTTCCAGCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGCGAGCCAGTGAAGC 571
QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180
Db 572 CGTGATTTATGTCAATGCAAAATGGGATTTGAAGACTGTGATGATGATCAGCCACCTAAGAA 631
QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGlnArgGluAla 200
Db 632 AAGAAATCCAGTTCAGCAAGAAAGAAACGCTCCAGGCGCAAGCAGGAGGAGGAGCT 691
QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db 692 TCACCTGTTGAGTTTGCATAGATCTTAAATGAACCTTACATCTGCTTATGCAACCAAGTG 751
QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGlnGlnCysProIleGluThrPheHisPhe 240
Db 752 TCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCCAATTTGAATGGTTTCACTTT 811
QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260
Db 812 TCATGTGTTTCACTTACCTTATAAACCAAGGAGGAAATGGTATTTGCCCAAGTGCAGGGA 871
```

QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280  
|||  
Db 872 GATAATGAGAAACAATGCAAAAGTACTGAAAGACAAAGGATAGATGAGG 931  
|||

## RESULT 3

US-09-484-970B-81  
; Sequence 81, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkumth, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 81  
; LENGTH: 1154  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6426186 197886.1CB1  
; NAME/KEY: unsure  
; LOCATION: 17-37  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-484-970B-81

## Alignment Scores:

Pred. No.: 2,64e-139 Length: 1154  
Score: 1334.00 Matches: 270  
Percent Similarity: 81.87% Conservative: 1  
Best Local Similarity: 81.57% Mismatches: 9  
Query Match: 90.07% Indels: 53  
DB: 4 Gaps: 1

US-09-513-365A-1 (1-280) x US-09-484-970B-81 (1-1154)

QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20  
|||  
Db 9 ATGTTAGNNNNNNNNNNNNNNNNNNNNNNCTGCGCTCGCTCTCGACCGGGAGCG 68  
|||  
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40  
|||  
Db 69 AC-CGGCTGCTACCTGCTACGTCGAGGACTCTTGAGTGGCTGGAGTGGCTGCCAC 127  
|||  
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln----- 57  
|||  
Db 128 GACATGCAGAGGAACGTCGTCTGTGCTGCGAGA-CTGGACAAACAATATCAAGTCCGGGC 186  
|||  
QY 57 ----- 57  
|||  
Db 187 TGTGCGCGGGCGTGTGTTGCGGCCCGAGCGAGTCCGAATCGGGTTTGAGCATGTTT 246  
|||  
QY 57 ----- 57  
|||  
Db 247 GCGGTGATGTTTCCAACTCTTTCCAGTCAATGATGATGATGATGATGATGATGATGAT 306  
|||  
QY 58 -----GluThrLeuLysGluLeuAspValTyrGluLysTyr 70  
|||  
Db 307 ACCGTGCGGATCTGCTCGCAACGTTTAAAGGAATGATGATGATGATGATGATGATGAT 366  
|||  
QY 70 rLysLysGluAspAspLeuAsnGlnLysLysArgLeuGlnGlnLeuGlnArgAlaLe 90  
|||  
Db 367 TAAGAAAGAGATGATTTAAACCAAGAAAGCTCTACAGCAGCTTCTCCAGAGACACT 426  
|||  
QY 90 uLeuAsnSerGlnGluLeuGlyAspGluLysLysLeuGlnLeuValThrGlnMetLeuGluLe 110  
|||  
Db 427 AATTAATAGTCAAGATTTGGAGATGAAAAAATACAGATTGTTTACACAAATGCTCAAT 486  
|||  
QY 110 uValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaG 130  
|||

Db 487 GGTGGAATTCGGGCAAGACAAATGAGTTACACTCACAGTGTTCCTCAAGATCCTGCTGA 546  
|||  
QY 130 uSerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerAr 150  
|||  
Db 547 AAGTGAACGAGCCTCAGATAAAGCAAGATGATTCAGCCCAACCAAGAAAGATCTTCAAG 606  
|||  
QY 150 gArgProArgArgGlnArgThrSerGluSerArgAspLeuCysHisMetAlaAsnGly 170  
|||  
Db 607 AAGACCCCGCAGCAGCGGACCGAGTGAAGCCGTGATTATGTCACATGGCAATGGGAT 666  
|||  
QY 170 eGluAspCysAspAspGlnProProLysGluLysSerLysSerLysSerAlaLysLys 190  
|||  
Db 667 TGAAGACTGTGATGATCAGCCACCTAAAGAAAAAGAAATCCAAGTCAGCAAGAAAAAGAA 726  
|||  
QY 190 sArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaLeuAspProAs 210  
|||  
Db 727 AGCTCCAAAGGCCAAGCAGGAAGGAAAGCTTCACCTGTTGAGTTTGCAATAGATCTCTAA 786  
|||  
QY 210 nGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAs 230  
|||  
Db 787 TGAACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTGACAA 846  
|||  
QY 230 nGluGlnCysProLysGluTyrPheHisPheSerCysValSerLeuThrTyrLysProLy 250  
|||  
Db 847 TGAACAGTGTCCAATGAAATGCTTTCACCTTTTCATGTTTCTACTTACCTATAAACCAAA 906  
|||  
QY 250 sGlyLysThrTyrCysProLysCysArgGlyAspAsnGluLysThrMetAspLysSerTh 270  
|||  
Db 907 GGGGAAATGTTATTCGCCAAAGTGCAGGGGAGATATATGAGAAACAATGACAAAAAGTAC 966  
|||

## RESULT 4

US-09-006-783A-4  
; Sequence 4, Application US/09006783A  
; Patent No. 6257366  
; GENERAL INFORMATION:  
; APPLICANT: Guckov, Andrey V  
; APPLICANT: Garkavstev, Igor  
; APPLICANT: Riabowol, Karl  
; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
; TITLE OF INVENTION: Pathway  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/006,783A  
; FILING DATE: 15-JAN-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6297366nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 97,837  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 873 base pairs  
; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..813
US-09-006-783A-4

```

## Alignment Scores:

```

Pred. No.: 6,3e-86 Length: 873
Score: 853.50 Matches: 161
Percent Similarity: 75.74% Conservative: 45
Best Local Similarity: 59.19% Mismatches: 51
Query Match: 57.63% Indels: 15
DB: 3 Gaps: 3

```

US-09-513-365A-1 (1-280) x US-09-006-783A-4 (1-873)

```

QY 18 GlyGluArgSerArgLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37
DB 25 GGGAGCAGCTCCACCTGTGTGAC---TATGTGGAGGACTACTGTGACTCCATCGAGTCC 81
QY 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln 57
DB 82 CTGCCTTTCGACTTGCAGAGAAATGTCGCTGATCGCGGAGATCGACGCGAAATACCAA 141
QY 58 GluThrLeuLysGluLeuAspValTyrGluLysTyrLysGluAspLeuAsn 77
DB 142 GAGATCCTGAAGAGCTAGACGAGTCTACGAGCGCTTCAGTCGCGAGACACAGCGGGCG 201
QY 78 GlnLysLysArgLeuGlnGlnLeuLeuArgAlaLeuLeuAsnSerGlnGluLeuGly 97
DB 202 CAGAAGCGCGGATGCTGACTGTGTGACGCGCGCTGATCCGACCCAGGAGCTGGGC 261
QY 98 AspGluLysLeuValThrGlnMetLeuLeuValGluLeuValGluLeuValGluLeu 117
DB 262 GACGAGAGATCCAGATCTGTGAGCCAGATGTTGTGAGATGTTGTGAGATGTTGTGAG 321
QY 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
DB 322 GTGACAGCCACCGTGGAGCTGTTGAGGCGCGACGAGGAGCTGGGCGACACAGTGGGCAAC 381
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnPro 145
DB 382 AGCGGCAAGGTGGCGCGGACAGCCCAATGGCGATGCGGTAGCGAGTCTGACAGCC 441
QY 146 GluArgSerSerArgArgProArgArgGlnArgThrSerGluSerArgAspLeuCysHis 165
DB 442 -----AACAGCAAGCGCTCAGCGGCGGCGGCAACACAGAGAACCGTGAGAACGGCTCC 495
QY 166 MetAlaAsnGlyLeuGluAspCysAspAspGlnProProLysGluLysLysSerLys 185
DB 496 AGCAACACGACGACGACGACGCGCGCTCGGGGACACCCCAAGGAGAGAGGCGCAAGCC 555
QY 186 AlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPhe 205
DB 556 TCCAGAGAGAGAGCGCTCCAGGCGCAAGCGGCGGAGAGCGCTCCCTCGCCGACCTC 615
QY 206 AlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMet 225
DB 616 CCCATCGACCCCAACGACCCAGCTACTGCTGTGTGCAACAGGTCTCTCTATGGGAGATG 675
QY 226 IleGlyCysAspAsnGluGlnCysProIleGluTyrPheHisPheSerCysValSerLeu 245
DB 676 ATCGGTGGGACCAACGACGAGTCCCATCGAGTGTTCACCTTCCTGCTGCGTGGGCGCTC 735
QY 246 ThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGluLysThr 265
DB 736 ATCATTAACCCGAGGCGAGTGTGTCTCCAGTGTCCGCGGGGAGAGAACGAGAGACC 795
QY 266 MetAspLysSerThrGluLysThrLysLysAspArg 277
DB 796 ATGGACAAAGCCCTCGGAGAAATCCAAAAAGAGAGG 831

```

## RESULT 5

```

US-09-258-257-1
; Sequence 1, Application US/09258257
; Patent No. 5965398
; GENERAL INFORMATION:
; APPLICANT: GARKAVTSEV, Igor
; APPLICANT: RIABOWOL, Karl
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A TUMOR
; TITLE OF INVENTION: SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/258,257
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/569,721
; FILING DATE: 08-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mooi, Leellie A.
; REGISTRATION NUMBER: 37,047
; REFERENCE/DOCKET NUMBER: 028722-128
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 854-7400
; TELEFAX: (650) 854-8275
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1902 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 109..738
US-09-258-257-1

```

## Alignment Scores:

```

Pred. No.: 1,94e-71 Length: 1902
Score: 728.00 Matches: 136
Percent Similarity: 74.79% Conservative: 39
Best Local Similarity: 58.12% Mismatches: 45
Query Match: 49.16% Indels: 14
DB: 2 Gaps: 2

```

US-09-513-365A-1 (1-280) x US-09-258-257-1 (1-1902)

```

QY 56 TyrGlnGluThrLeuLysGluIleAspValTyrGluLysTyrLysGluAspAsp 75
DB 31 TGGAAACAGACTCTGAGGAGCTAGCAGGTCTACGAGCGCTTCAGTCGCGAGACAGAC 90
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlu 95
DB 91 GGGGCGGACAGAGCGGCGATGCTGCTGTGTGCGCGCGCTGATCCGCGAGAGAG 150
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluLeuArgAla 115
DB 151 CTGGGCGGAGAGAAATCCAGATCTGAGCCAGATGTTGTGAGTGTGTGAGAACCCGACG 210
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130

```

Db 211 CGCAGGTGGACAGCCACAGTGTGGAGCTTTCGAGCGCGCAGCAGGAGTGGCGGCACACAGTG 270  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
Db 271 GGCAACAGCGGCAAGGTTGGCGGACAGCCCAATGGCGATGGGTAGCGCAGTCTGAC 330  
QY 144 GlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSerArgAspLeu 163  
Db 331 AAGCCC-----AACAGCAAGCGCTCACGGCGGCGAGCAACACGAGAACCGTGAGAAC 384  
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSer 183  
Db 385 GCGTCAGCAACACGACGACGAGCGCGCTCGGCGCACCCCAAGGAGAGAGGCC 444  
QY 184 LysSerAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 203  
Db 445 AAGACCTCCCAAGAGAGAGCGCTCCAGGCCCAAGCGCGAGGAGGCGTCCCTGCC 504  
QY 204 GluPheAlaIleAspProAsnGluProThrTyxCysLeuCysAsnGlnValSerTyrgly 223  
Db 505 GACCTCCCATCGACCCCAACGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 564  
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisSerCysVal 243  
Db 565 GAGATGATCGCTGCGGACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 624  
QY 244 SerLeuThrTyrglyProLysLysLysLysLysLysLysLysLysLysLysLysLys 263  
Db 625 GGCTCAATCATATAACCCAGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 684  
QY 264 LysThrMetAspLysSerThrGluLysLysLysLysLysLysLysLysLysLysLys 277  
Db 685 AAGACCATGACAAAGCCCTGGAGAAATCCAAAAAGAGAGG 726

## RESULT 6

US-09-258-371-1  
; Sequence 1, Application US/09258371  
; Patent No. 5986078  
; GENERAL INFORMATION:  
; APPLICANT: Garkavtsev, Igor  
; APPLICANT: Ribowol, Karl  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
; TITLE OF INVENTION: SUPPRESSOR GENE INGI  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/258,371  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/751,230  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mooi, Leslie A.  
; REGISTRATION NUMBER: 37,047  
; REFERENCE/DOCKET NUMBER: 028722-144  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-7400  
; TELEFAX: 415-854-8275  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1902 base pairs

; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 109..741  
; US-09-258-371-1

## Alignment Scores:

Prod. No.: 1,94e-71 Length: 1902  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
DB: 2 Gaps: 2

US-09-513-365A-1 (1-280) x US-09-258-371-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluLysAspValTyrglyLysTyrglyLysGluAsp 75  
Db 31 TGGAAACAGATCTCTGAAGGAGCTAGACGAGTGTACGAGCGCTTCAGTCGCGAGACAGAC 90  
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlu 95  
Db 91 GGGCGCGCAGAAAGCGCGGATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 150  
QY 96 LeuGlyAspGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 115  
Db 151 CTGGGCGATCAGAAAGATCCAGATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 210  
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
Db 211 CGCAGGTGGACAGCCACGCTGT 270  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSer 143  
Db 271 GGCAACAGCGGCAAGGTTGGCGGACAGCCCAATGGCGATGGGTAGCGCAGTCTGAC 330  
QY 144 GlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSerArgAspLeu 163  
Db 331 AAGCCC-----AACAGCAAGCGCTCACGGCGGCGAGCAACACGAGAACCGTGAGAAC 384  
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSer 183  
Db 385 GCGTCAGCAACACGACGACGAGCGCGCTCGGCGCACCCCAAGGAGAGAGGCC 444  
QY 184 LysSerAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 203  
Db 445 AAGACCTCCCAAGAGAGAGCGCTCCAGGCCCAAGCGCGAGGAGGCGTCCCTGCC 504  
QY 204 GluPheAlaIleAspProAsnGluProThrTyrglyLysLysLysLysLysLysLysLys 223  
Db 505 GACCTCCCATCGACCCCAACGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 564  
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisSerCysVal 243  
Db 565 GAGATGATCGCTGCGGACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 624  
QY 244 SerLeuThrTyrglyProLysLysLysLysLysLysLysLysLysLysLysLysLys 263  
Db 625 GGCTCAATCATATAACCCAGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 684  
QY 264 LysThrMetAspLysSerThrGluLysLysLysLysLysLysLysLysLysLysLys 277  
Db 685 AAGACCATGACAAAGCCCTGGAGAAATCCAAAAAGAGAGG 726

## RESULT 7

US-08-569-721A-1  
; Sequence 1, Application US/08569721A  
; Patent No. 6037121  
; GENERAL INFORMATION:  
; APPLICANT: GARKAVTSEV, Igor

APPLICANT: RIABOWOL, Karl  
 TITLE OF INVENTION: DNA SEQUENCE ENCODING A TUMOR  
 TITLE OF INVENTION: SUPPRESSOR GENE  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/569,721A  
 FILING DATE: 08-DEC-1995  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mooi, Leslie A.  
 REGISTRATION NUMBER: 37,047  
 REFERENCE/DOCKET NUMBER: 028722-128  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 854-7400  
 TELEFAX: (650) 854-8275  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1902 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 109..738  
 US-08-569-721A-1

## Alignment Scores:

Pred. No.: 1.94e-71 Length: 1902  
 Score: 728.00 Matches: 136  
 Percent Similarity: 74.79% Conservative: 39  
 Best Local Similarity: 58.12% Mismatches: 45  
 Query Match: 49.16% Indels: 14  
 DB: 3 Gaps: 2

US-09-513-365A-1 (1-280) x US-08-569-721A-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspApp 75  
 Db 31 TGGAAACACATCTTGAAGAGCTAGACGAGTGTCTACGAGCGCTTCAGTCGCGAGACAG 90  
 QY 76 LeuAsnGlnLysLysArgLeuGlnLeuGlnAlaLeuLeuLeuLeuLeuLeuLeuLeu 95  
 Db 91 GGGGCCACAGCGCGGAGTGTGCACTGTGTGACGCGCGCTGTATCGCAGCCAGGAG 150  
 QY 96 LeuGlyAspGluLysLeuGlnLeuValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
 Db 151 CTGGCGCAGCAGAGAATCCAGATCGTGTGAGCAGATGTGTGAGTGTGTGAGAGAACCGCAG 210  
 QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
 Db 211 CGCAGGTGGACAGCCAGCTGTGAGAGTGTGTGAGAGCCAGCAGGAGTGGCGCAGCAGTG 270  
 QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSer 143  
 Db 271 GGCACAGCGCGGAGGTTGGCGGAGCAGGCCAATGGCGATGCGGTAGCGAGTCTGAC 330  
 QY 144 GlnProGluArgSerArgProArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArg 163  
 Db 331 AAGCCC-----AACAGCAACGCGCTCACGGCGCAGCGCAACCAACGAGAACCGTGAGAAC 384

QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspGlnProProLysGluLysLysSer 183  
 Db 385 GCGTCCAGCAACACGACGACGCGCGCTCGGCGCACACCCCAAGGAGAGAGGCC 444  
 QY 184 LysSerAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 203  
 Db 445 AAGACCTCCAAG 504  
 QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
 Db 505 GACCTCCCATCGACGACCCCAACGACCCACGACTGTGTGTGCAACAGGTCTCTATGG 564  
 QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243  
 Db 565 GAGATGATCGCTGCGACACGACGAGTGGCCCATCGAGTGGTTCACCTTCTCGTGGCG 624  
 QY 244 SerLeuThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspGlu 263  
 Db 625 GGGCTCAATCATAAACCCCAAGGCAAGTGTACTGTCCCAAGTCCCGGGGAGAACGAG 684  
 QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
 Db 685 AAGACCATGGACAAAGCCCTGGAGAAATCCAAAGAGAGAG 726

## RESULT 8

US-08-751-230-1  
 ; Sequence 1, Application US/08751230  
 ; Patent No. 6117633  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garkavtsev, Igor  
 ; APPLICANT: Riabowol, Karl  
 ; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
 ; TITLE OF INVENTION: SUPPRESSOR GENE INGI  
 ; NUMBER OF SEQUENCES: 23  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis  
 ; STREET: 699 Prince Street  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-1404  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/751,230  
 ; FILING DATE: 15-NOV-1996  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/569721  
 ; FILING DATE: 08-DEC-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mooi, Leslie A.  
 ; REGISTRATION NUMBER: 37,047  
 ; REFERENCE/DOCKET NUMBER: 028722-144  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-854-7400  
 ; TELEFAX: 415-854-8275  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1902 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 109..741  
 ; US-08-751-230-1  
 Alignment Scores:

```
Pred. No.: 1,94e-71 Length: 1902
Score: 728.00 Matches: 136
Percent Similarity: 74.79% Conservative: 39
Best Local Similarity: 58.12% Mismatches: 45
Query Match: 49.16% Indels: 14
DB: 3 Gaps: 2

US-09-513-365A-1 (1-280) x US-08-751-230-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspAsp 75
Db 31 TGGAAACAGATCTGAGAGAGCTAGACGAGTCTACGAGCGCTTACGTCCGAGACAGAC 90
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuLeuSerGlnGlu 95
Db 91 GGGGCGCAGAGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115
Db 151 CTGGGCGCAGAGAGATCCAGATCTGAGCAGATGCTGAGTGTGGAGAACCGCAGC 210
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
Db 211 CGGAGGTGGACAGCCAGCTGGAGGCTTTCGAGGCGCAGCAGAGCTGGGCGCAGACAG 270
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143
Db 271 GGCAACAGCGCAAGGTTGGCGCGACAGGCCCAATGGCGTCCGAGCGAGTCTGAC 330
QY 144 GlnProGluArgSerSerArgArgProArgGlnArgThrSerGluSerArgAspLeu 163
Db 331 AAGCCC-----AACAGCAAGCGCTCAGCGCGCAGCGCAACAAACAGAGAACCGTGAGA 384
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysSer 183
Db 385 GGTCCAGCAACACAGCAGCAGCGCGCTCGGCGCACCCCAAGAGAGAGAGGCC 444
QY 184 LysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203
Db 445 AGACCTCCAGAGAGAGAGCGCTCCAGGCCAAGCGCGGAGCGGAGCGTCCCTGCC 504
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223
Db 505 GACCTCCCATCCAGCCCAACGAGACCCAGCTACTGTCTGTGCAACCGAGTCTCTATGG 564
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243
Db 565 GAGATGATCGGCTGGCAGCAACAGCAGAGTCCCATCGAGTGTTCCTCTCTGTCGTG 624
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263
Db 625 GGGCTCAATCATTAACCCAGGCGCAAGTGTACTGTCCCAAGTCCCGGGGGAGAACGAG 684
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277
Db 685 AAGACCATGCAAAAGCCCTGGAGAAATCCAAAAAGAGAGG 726
```

## RESULT 9

```
US-09-499-082-1
Sequence 1, Application US/09499082
Patent No. 6143522
GENERAL INFORMATION:
APPLICANT: Helbing, Caren C.
APPLICANT: Riabowol, Karl
APPLICANT: Johnston, Randall N.
APPLICANT: Garkavtsev, Igor
TITLE OF INVENTION: METHODS OF MODULATING APOPTOSIS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
```

```
COUNTRY: USA
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/499,082
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/828,158
FILING DATE: 27-MAR-1997
APPLICATION NUMBER: US 08/751230
FILING DATE: 15-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/569721
FILING DATE: 08-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Mooi, Leslie A.
REGISTRATION NUMBER: 37,047
REFERENCE/DOCKET NUMBER: 028722-148
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-854-7400
TELEFAX: 650-854-8275
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1902 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 109..741
US-09-499-082-1

Alignment Scores:
Pred. No.: 1,94e-71 Length: 1902
Score: 728.00 Matches: 136
Percent Similarity: 74.79% Conservative: 39
Best Local Similarity: 58.12% Mismatches: 45
Query Match: 49.16% Indels: 14
DB: 3 Gaps: 2

US-09-513-365A-1 (1-280) x US-09-499-082-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspAsp 75
Db 31 TGGAAACAGATCTGAGAGAGCTAGACGAGTCTACGAGCGCTTACGTCCGAGACAGAC 90
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuLeuSerGlnGlu 95
Db 91 GGGGCGCAGAGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115
Db 151 CTGGGCGCAGAGAGATCCAGATCTGAGCAGATGCTGAGTGTGGAGAACCGCAGC 210
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
Db 211 CGGAGGTGGACAGCCAGCTGGAGGCTTTCGAGGCGCAGCAGAGCTGGGCGCAGACAG 270
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143
Db 271 GGCAACAGCGCAAGGTTGGCGCGACAGGCCCAATGGCGTCCGAGCGAGTCTGAC 330
QY 144 GlnProGluArgSerSerArgArgProArgGlnArgThrSerGluSerArgAspLeu 163
Db 331 AAGCCC-----AACAGCAAGCGCTCAGCGCGCAGCGCAACAAACAGAGAACCGTGAGA 384
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysSer 183
```







```
QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAsp 75
Db 190 TGGAAACAGATCTTGAAGAGCTAGACGAGTCTACGAGCGCTTACGTCGGACAGAC 249
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnAlaLeuLeuLeuSerGlnGlu 95
Db 250 GGGGGCCAGAGCGGGGATGTCACCTGTGTCCAGCGCGCTGTATCCGACGACGAG 309
QY 96 LeuGlyAspGluLysLeuGlnLeuValThrGlnMetLeuLeuValGluLeuArgAla 115
Db 310 CTGGGGCCAGAGAGATCCAGATCGTGTGAGCCAGATGTGTGTGTGTGTGTGTGTGT 369
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
Db 370 CGCGAGGTGGACAGCCAGCGGAGCTGTTCGAGCGCGACGAGGATGGGCGACACAG 429
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143
Db 430 GGCACAGCGCGCAAGGTTGGCGCGACAGCCCAATGGCGATCGGTAGCGCAGTCTGAC 489
QY 144 GlnProGluArgSerArgProArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 163
Db 490 AAGCCC-----AACAGCAAGCGCTCACGCGCGCAGCGCAACACGAGAACCGTGAGA 543
QY 164 CysHisMetAlaAsnGlyLeuGluAspCysAspGlnProProlLysGluLysSer 183
Db 544 GCGTCCAGCAACCCAGCCAGCGCGCGCTCGGCGCACACCCCAAGGAGAGAGGCC 603
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGlnGlnGlnGlnGlnGlnGln 203
Db 604 AAGACTCCAGAGAGAGAGAGCGCTCCAGGCCAGCGGCGGAGCGAGGCGTCCCTGCC 663
QY 204 GluPheAlaLeuAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223
Db 664 GACCTCCCATCAGCCCAACGAGCAACCCAGCTACTGTGTGTGTGTGTGTGTGTGTGT 723
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluThrPheHisPheSerCysVal 243
Db 724 GAGATGATCGGTGGCGACCAACGAGCGTGGCGCGCTCGATCGAGTGTTCACCTTCTCGT 783
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263
Db 784 GGGCTCAATCATTAACCCAGGCGCAAGTGTACTGTCTCCCAAGTCCCGGGGGAGAACGAG 843
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277
Db 844 AAGACCATGACAAAGCCCTGGAGAAATCCAAAAAGAGAGG 885
```

## RESULT 13

```
US-09-499-082-9
; Sequence 9, Application US/09499082
; Patent No. 6143522
; GENERAL INFORMATION:
; APPLICANT: Helbing, Caren C.
; APPLICANT: Ribowol, Karl
; APPLICANT: Johnston, Randall N.
; APPLICANT: Garkavtsev, Igor
; TITLE OF INVENTION: METHODS OF MODULATING APOPTOSIS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/499,082
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,158
; FILING DATE: 27-MAR-1997
; APPLICATION NUMBER: US 08/751230
; FILING DATE: 15-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/569721
; FILING DATE: 08-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mooi, Leslie A.
; REGISTRATION NUMBER: 37,047
; REFERENCE/DOCKET NUMBER: 028722-148
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-854-7400
; TELEFAX: 650-854-8275
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..900
; US-09-499-082-9
```

```
Alignment Scores:
Pred. No.: 2,188-71 Length: 2061
Score: 728.00 Matches: 136
Percent Similarity: 74.79% Conservative: 39
Best Local Similarity: 58.12% Mismatches: 45
Query Match: 49.16% Indels: 14
DB: 3 Gaps: 2
```

US-09-513-365A-1 (1-280) x US-09-499-082-9 (1-2061)

```
QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAsp 75
Db 190 TGGAAACAGATCTTGAAGAGCTAGACGAGTCTACGAGCGCTTACGTCGGACAGAC 249
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnAlaLeuLeuLeuSerGlnGlu 95
Db 250 GGGGGCCAGAGCGGGGATGTCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 309
QY 96 LeuGlyAspGluLysLeuGlnLeuValThrGlnMetLeuLeuValGluLeuArgAla 115
Db 310 CTGGGGCCAGAGAGATCCAGATCGTGTGAGCGCGACGAGGATGGGCGACACAG 369
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
Db 370 CGCGAGGTGGACAGCCAGCGGAGCTGTTCGAGCGCGACGAGGATGGGCGACACAG 429
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143
Db 430 GGCACAGCGCGCAAGGTTGGCGCGACAGCCCAATGGCGATCGGTAGCGCAGTCTGAC 489
QY 144 GlnProGluArgSerArgProArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 163
Db 490 AAGCCC-----AACAGCAAGCGCTCACGCGCGCAGCGCAACACGAGAACCGTGAGA 543
QY 164 CysHisMetAlaAsnGlyLeuGluAspCysAspGlnProProlLysGluLysSer 183
Db 544 GCGTCCAGCAACCCAGCCAGCGCGCGCTCGGCGCACACCCCAAGGAGAGAGGCC 603
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGlnGlnGlnGlnGlnGlnGln 203
Db 604 AAGACTCCAGAGAGAGAGAGCGCTCCAGGCCAGCGGCGGAGCGAGGCGTCCCTGCC 663
QY 204 GluPheAlaLeuAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223
```

Db 664 GACCTCCCATCAGCCACCAACCAACAGTACTGTGTGCAACAGGTCTCTATGGG 723  
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrPheHisPheSerCysVal 243  
Db 724 GAGATGATCGGTGCGCAACAGCAGAGTGTGCGCCATCGAGTGTTCACCTTCTCTGTCGGTG 783  
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263  
Db 784 GGGCTCAATCAATAAACCAGGCAAGTGTGTCTGCTCCCAAGTGTCCCGGGGAGAACGAG 843  
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 844 AAGACCATGACAAAGCCCTGGAGNAATCCAAAAAGAGAGG 885

## RESULT 14

US-09-258-372-9  
; Sequence 9, Application US/09258372  
; Patent No. 6238918  
; GENERAL INFORMATION:  
; APPLICANT: Garkavtsev, Igor  
; APPLICANT: Riabowol, Karl  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
; TITLE OF INVENTION: SUPPRESSOR GENE INGI  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/258,372  
; FILING DATE:

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/751,230  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mooi, Leelie A.  
; REGISTRATION NUMBER: 37,047  
; REFERENCE/DOCKET NUMBER: 028722-144  
; TELEPHONE: 415-854-7400  
; TELEFAX: 415-854-8275

INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2061 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 16..900

US-09-258-372-9

Alignment Scores:

Pred. No.:	2,18e-71	Length:	2061
Score:	728.00	Matches:	136
Percent Similarity:	74.79%	Conservative:	39
Best Local Similarity:	58.12%	Mismatches:	45
Query Match:	49.16%	Indels:	14
DB:	3	Gaps:	2

US-09-513-365A-1 (1-280) x US-09-258-372-9 (1-2061)

QY 56 TyrGlnGluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspAsp 75  
Db 190 TGGAAACAGATCTCTAAGGAGCTAGCAGAGTGTCTACGAGCGCTTCACTCGCAGACAGAC 249  
QY 76 LeuAsnGluLysArgLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 95  
Db 250 GGGGCGCAGAACCGCGGATGCTGTCAGTGTGCGAGCGCGCTGATCCGCGACGAGGAG 309  
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuLeuLeuValGluLeuAsnArgAla 115  
Db 310 CTGGGCGCAGCAAGATCCAGATCTGAGCAGAGTGTGAGAGTGTGTGAGAGACCGCAGC 369  
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu 130  
Db 370 CGGCGAGTGGACAGCCAGCGTGTTCGAGCGCGCAGCAGAGTGTGGCGCAGACAGTG 429  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
Db 430 GGCACACGCGCAAGGTTGGCGCGCAGCAGCCCAATGGCGATGCGGTAGCGCAGTCTGAC 489  
QY 144 GlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSerArgAspLeu 163  
Db 490 AAGCCC-----AACAGCAAGCGCTCACGCGCGCAGCGCAACACGAGAACCGTGAGAAC 543  
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSer 183  
Db 544 GCGTCCAGCAACACGACGACGACGCGCGCTCGGGGCACACCCCAAGGAGGAGGCC 603  
QY 184 LysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
Db 604 AAGACCTCCAGAGAAGAGCGCTCCAGGCGCAAGCGCGAGCGAGAGCGCTCCCTTGCC 663  
QY 204 GluPheAlaLeuAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
Db 664 GACCTCCCATCGACCCCAACGACCCAGTGTGTGTGCAACAGGTCTCTATGGG 723  
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrPheHisPheSerCysVal 243  
Db 724 GAGATGATCGGTGCGCAGCAACGAGTGTGCGCCATCGAGTGTTCACCTTCTCTGCGTG 783  
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263  
Db 784 GGGCTCAATCAATAAACCAGGCAAGTGTGTCTCCCAAGTGTCCCGGGGAGAACGAG 843  
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 844 AAGACCATGACAAAGCCCTGGAGNAATCCAAAAAGAGAGG 885

## RESULT 15

US-09-159-871-1  
; Sequence 1, Application US/09159871A  
; Patent No. 6420136  
; GENERAL INFORMATION:  
; APPLICANT: RIABOWOL, Karl T.  
; TITLE OF INVENTION: METHOD OF MODULATING P53 ACTIVITY  
; FILE REFERENCE: 028722-181  
; CURRENT APPLICATION NUMBER: US/09/159,871A  
; CURRENT FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: US 60/060,138  
; PRIOR FILING DATE: 1997-09-26  
; PRIOR APPLICATION NUMBER: US 09/006,783  
; PRIOR FILING DATE: 1998-01-14  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2061  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (16)..(897)  
US-09-159-871-1

## Alignment Scores:

Pred. No.:	2,18e-71	Length:	2061
Score:	728.00	Matches:	136
Percent Similarity:	74.79%	Conservative:	39
Best Local Similarity:	58.12%	Mismatches:	45
Query Match:	49.16%	Indels:	14
DB:	4	Gaps:	2

US-09-513-365A-1 (1-280) x US-09-159-871-1 (1-2061)

```
QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspAsp 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 TGAACACAGATCTGAAGAGAGTAGACGAGTGTACGAGCGCTTACGTCGAGACAGAC 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuSerGlnGlu 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 GGGGCGCAGAAGCGCGGATGTGTCACGTGTGTCAGCGCGCTGATCCGACGCCAGGAG 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 CTGGGCGACGAGNAGATCAGATCGTAGCCAGATGGTGGAGCTGGTGGAGAACCGCAGC 369
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 CGGCAGGTGCACAGCCACGTCGGAGCTGTCGAGCGCAGCAGGAGCTGGCGCACAGTG 429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 430 GGCACAGCGCAAGGTTGCGCGGACAGGCCCAATGCGCGATGCGGATCGCGAGTCTGAC 489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 GlnProGluArgSerSerArgProArgArgGlnArgThrSerGluSerArgAspLeu 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 AAGCCC-----AACAGCAGAGCGCTCACGGCGGCAGCGCAACACGAGAACCGTGAGAAC 543
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSer 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 544 GCGTCCAGCAACACACGACCACGACGCGCGCTCGGGCACACCCCAAGGAGAGAGGCC 603
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 LysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 604 AAGACCTCCAGAAGAAGAAGCGCTCCAAGGCGAAGCGGAGGAGGAGCGTCCCTGCC 663
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 664 GACCTCCCATCGACCCCAACGAAACCCACGTAAGTGTGTGTGCAACGAGGTCTCTATGGG 723
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 724 GAGATGATCGGCTGCACAAACGACGAGTGCCTCCATCGAGTGGTTCCTTCTCGTGGTG 783
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 784 GGGCTCAATCATAAACCCCAAGGCGAAGTGTACTGTCCCAAGTGGCGGGGAGAACGAG 843
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 844 AAGACCATGGACAAAGCCCTGGAGAAATCCAAAAAGAGAG 885
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: January 12, 2004, 19:49:53

Job time : 84 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 17:22:12 ; Search time 430 Seconds  
(without alignments)

```

Title: US-09-513-365A-2
Perfect score: 1080
Sequence: 1 gcggccgcggccggtgatg.....aaaaaaaaaaaaaaaaaaaa 1080

```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues

Total number of hits satisfying chosen parameters: 4552328

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:2:

14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*

17: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*

18: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1046	96.9	1078	14	US-10-115-899-7		Sequence 7, Appli
2	838.4	77.6	840	14	US-10-115-899-6		Sequence 6, Appli
3	378.4	35.0	421	11	US-09-918-995-37033		Sequence 37033, A
4	237.4	22.0	873	11	US-09-968-653A-4		Sequence 4, Appli
5	183.4	17.9	8487	10	US-09-764-877-3454		Sequence 3454, Ap
6	182.2	16.9	2061	11	US-09-968-653A-2		Sequence 2, Appli
7	174.4	16.1	633	11	US-09-968-653A-6		Sequence 6, Appli
8	91.6	8.5	1864	13	US-10-117-720-435		Sequence 435, App
9	91.6	8.5	1864	13	US-10-117-720-435		Sequence 435, App
10	83.6	7.7	451	10	US-09-738-973-146		Sequence 146, App
11	83.6	7.7	451	10	US-09-854-133-146		Sequence 146, App
12	83.6	7.7	451	15	US-10-144-649A-146		Sequence 146, App
13	66.6	6.2	993	10	US-09-801-368-257		Sequence 257, App
14	62	5.7	813	13	US-09-938-842A-1178		Sequence 1178, A
15	60	5.6	60	13	US-09-908-975-12517		Sequence 12517, A

16	58.6	5.4	65	13	US-09-908-975-23744	Sequence 23744, A
17	56.8	5.3	456	11	US-09-918-995-16047	Sequence 16047, A
18	54	5.0	897	13	US-10-032-585-6348	Sequence 6348, Ap
C 19	52.8	4.9	249	13	US-09-814-353-5815	Sequence 5815, Ap
C 20	52.8	4.9	249	13	US-09-814-353-12097	Sequence 12097, A
21	51.6	4.8	350	13	US-09-814-353-18240	Sequence 18240, A
C 22	51.6	4.8	393	10	US-09-960-352-4582	Sequence 4582, Ap
C 23	50.8	4.7	361	10	US-09-919-580-114	Sequence 114, Appl
C 24	49.6	4.6	3673778	13	US-10-312-841-2	Sequence 2, Appl
C 25	49.2	4.6	495	13	US-09-814-353-5212	Sequence 5212, Ap
C 26	49.2	4.6	495	13	US-09-814-353-11499	Sequence 11499, A
C 27	49	4.5	593	11	US-09-871-161-262	Sequence 262, Ap
C 28	49	4.5	6071	13	US-10-311-455-298	Sequence 298, Appl
C 29	48.8	4.5	277	10	US-09-960-352-12673	Sequence 12673, A
C 30	48.6	4.5	912	15	US-10-128-714-2502	Sequence 2502, Ap
31	48.6	4.5	1254	15	US-10-128-714-1502	Sequence 1502, Ap
32	48.6	4.5	1254	15	US-10-128-714-6502	Sequence 6502, Ap
33	48.6	4.5	1254	15	US-10-128-714-7502	Sequence 7502, Ap
34	48.6	4.5	3254	15	US-10-128-714-502	Sequence 502, Appl
35	48.6	4.5	3254	15	US-10-128-714-5502	Sequence 5502, Ap
36	48.4	4.5	5520	14	US-10-001-887-43	Sequence 43, Appl
C 37	48.4	4.5	6314	13	US-10-240-452-14	Sequence 14, Appl
C 38	48	4.4	527	15	US-10-106-698-863	Sequence 863, Appl
C 39	47.6	4.4	5689	13	US-10-240-453-100	Sequence 100, Appl
C 40	47.6	4.4	5689	15	US-10-239-676-90	Sequence 90, Appl
C 41	47.6	4.4	8805	13	US-10-311-455-486	Sequence 486, Appl
C 42	47.6	4.4	3673778	13	US-10-312-841-1	Sequence 1, Appl
C 43	47.2	4.4	624	15	US-10-066-543-491	Sequence 491, Appl
C 44	47	4.4	18011	13	US-10-311-455-7	Sequence 7, Appl
45	46.8	4.3	18154	13	US-10-311-455-227	Sequence 227, Appl

## ALIGNMENTS

RESULT 1  
US-10-115-899-7  
; Sequence 7, Application US/10115899  
; Publication No. US20020151025A1  
; GENERAL INFORMATION:  
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human INGIL gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/10/115,899  
; CURRENT FILING DATE: 2002-04-05  
; PRIORITY APPLICATION NUMBER: 09/601,478  
; PRIORITY FILING DATE: 2000-08-03  
; PRIORITY APPLICATION NUMBER: JP H10-134679  
; PRIORITY FILING DATE: 1998-04-28  
; PRIORITY APPLICATION NUMBER: JP H10-73234  
; PRIORITY FILING DATE: 1998-03-05  
; PRIORITY APPLICATION NUMBER: JP H10-38133  
; PRIORITY FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1078

OK

61 CGCAGGAGTGTAGGGCAGCAGCAGCAACTGTACTCGTCGGCTCGCTCCTGACCGG 120  
Db |||||||  
85 CGCAGGAGTGTAGGGCAGCAGCAGCAACTGTACTCGTCGGCTCGCTCCTGACCGG 144  
Qy 121 GGAGCGAGCGCGTCTCACTCTGCTACGTCAGGAGTACCTTTGAGTGGCTGAGTGGCT 180  
Db 145 GGAGCGAGCGCGTCTCACTCTGCTACGTCAGGAGTACCTTTGAGTGGCTGAGTGGCT 204  
Qy 181 GCCCCAGCATCGCAGAGGAACGTCTGTGCTGCGAGAGCTGGACACAAATATCAAGA 240  
Db 205 GCCCCAGCATCGCAGAGGAACGTCTGTGCTGCGAGAGCTGGACACAAATATCAAGA 264  
Qy 241 AACGTTAAAGGAATTCATGATGCTACGAAATATTAAGAAAGAGATGATTTAAACCA 300  
Db 265 AACGTTAAAGGAATTCATGATGCTACGAAATATTAAGAAAGAGATGATTTAAACCA 324  
Qy 301 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 360  
Db 325 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 384  
Qy 361 TGAATAATACAGATGTTACACAAATGCTGGAATTTGGTGGAAATCGGCCAAGACAAT 420  
Db 385 TGAATAATACAGATGTTACACAAATGCTGGAATTTGGTGGAAATCGGCCAAGACAAT 444  
Qy 421 GGAGTTACACTCACAGTGTTCACAAAGATCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 480  
Db 445 GGAGTTACACTCACAGTGTTCACAAAGATCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 504  
Qy 481 AAAGATGGATTTCAGGCCAACACAGAAAGATCTTCAAGAAAGCCCGCAGGAGCGGACCAG 540  
Db 505 AAAGATGGATTTCAGGCCAACACAGAAAGATCTTCAAGAAAGCCCGCAGGAGCGGACCAG 564  
Qy 541 TGAAGCGGTGATTTATGTACATGCGCAATGGGATTTGAAGATGTGATGATCAGCCACC 600  
Db 565 TGAAGCGGTGATTTATGTACATGCGCAATGGGATTTGAAGATGTGATGATCAGCCACC 624  
Qy 601 TAAAGAAAGAAATCAAGTCAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 660  
Db 625 TAAAGAAAGAAATCAAGTCAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 684  
Qy 661 GGAAGCTTCACTGTGTGAGTTTCAATAGATCTTAATGAACCTACATCTGCTTTATGCAA 720  
Db 685 GGAAGCTTCACTGTGTGAGTTTCAATAGATCTTAATGAACCTACATCTGCTTTATGCAA 744  
Qy 721 CCAAGTGTCTTATGGGAGATGATAGATGTGACATGAACAGTGTCCAATTTGAATGGTT 780  
Db 745 CCAAGTGTCTTATGGGAGATGATAGATGTGACATGAACAGTGTCCAATTTGAATGGTT 804  
Qy 781 TCACCTTTTCATGTGTTTCACTTACCTATTAACCAAGGGGAATGGTATTGCCCAAGTG 840  
Db 805 TCACCTTTTCATGTGTTTCACTTACCTATTAACCAAGGGGAATGGTATTGCCCAAGTG 864  
Qy 841 CAGGGAGATAATGAGAAACAAATGACAAAGTACTGAAAGACAAAGAAAGATAGAAG 900  
Db 865 CAGGGAGATAATGAGAAACAAATGACAAAGTACTGAAAGACAAAGAAAGATAGAAG 924  
Qy 901 ATCAGAGTAAAGGCAATCCACATTTTAAAGGTTATTTGCTTTTATATATTCGTT 960  
Db 925 ATCAGAGTAAAGGCAATCCACATTTTAAAGGTTATTTGCTTTTATATATTCGTT 984  
Qy 961 TGCTTTTCAGAAATGTTTATAGGCTAAATGATGATGATGATGATGATGATGATGAT 1020  
Db 985 TGCTTTTCAGAAATGTTTATAGGCTAAATGATGATGATGATGATGATGATGATGAT 1044  
1021 GTATTAAATGGTGTATTAAGTGTGTGACTTTG 1054  
1045 GTATTAAATGGTGTATTAAGTGTGTGACTTTG 1078

## RESULT 2

US-10-115-899-6

; Sequence 6, Application US/10115899

; Publication No. US20020151025A1

GENERAL INFORMATION:  
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human INGIL gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/10/115,899  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/601,478  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
US-10-115-899-6

Query Match 77.6%; Score 838.4; DB 14; Length 840;  
Best Local Similarity 99.9%; Pred. No. 2e-190; 1; Indels 0; Gaps 0;  
Matches 839; Conservative 0; Mismatches 0; Gaps 0;

Qy 68 ATGTAGGCGCAGCAGCAGCAACTGTACTCGTCGGCTGCGCTCCTGACCGGGAGCGG 127  
Db 1 ATGTAGGCGCAGCAGCAGCAACTGTACTCGTCGGCTGCGCTCCTGACCGGGAGCGG 60  
Qy 128 AGCGGCTGCTCACCTGCTACGTCGAGGACTACCTTGTGAGTGGAGTGGTGGCTGCTG 187  
Db 61 AGCGGCTGCTCACCTGCTACGTCGAGGACTACCTTGTGAGTGGAGTGGTGGCTGCTG 120  
Qy 188 GACATGCGAGGAACTGTGCTGTCGCGAGAGCTGGACAAACAAATATCAAGAAAGCTTA 247  
Db 121 GACATGCGAGGAACTGTGCTGTCGCGAGAGCTGGACAAACAAATATCAAGAAAGCTTA 180  
Qy 248 AAGGAAATTTGATGATGCTTACGAAATAATAAGAAAGAGATGATTTAAACCAAGAA 307  
Db 181 AAGGAAATTTGATGATGCTTACGAAATAATAAGAAAGAGATGATTTAAACCAAGAA 240  
Qy 308 CGTCTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGAGATGA 367  
Db 241 CGTCTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGAGATGA 300  
Qy 368 ATACAGATTTGTACACAAATGCTCGAAATTTGGTGGAAATCGGGCAAGACAAATGGAGTTA 427  
Db 301 ATACAGATTTGTACACAAATGCTCGAAATTTGGTGGAAATCGGGCAAGACAAATGGAGTTA 360  
Qy 428 CACTCACAGTGTTCACAGATCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATG 487  
Db 361 CACTCACAGTGTTCACAGATCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATG 420  
Qy 488 GATTCAGCAACCAAGAAAGATCTTCAAGAAAGCCCGCAGCAGCGGACCAAGTGAAGC 547  
Db 421 GATTCAGCAACCAAGAAAGATCTTCAAGAAAGCCCGCAGCAGCGGACCAAGTGAAGC 480  
Qy 548 CGTGATTTATGTACATGCAATGCGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAA 607  
Db 481 CGTGATTTATGTACATGCAATGCGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAA 540  
Qy 608 AAGAAATCCAGTTCAGCAAGAAAGAAACGCTCCAGGCCCAAGCAGGAAAGGGAGCT 667  
Db 541 AAGAAATCCAGTTCAGCAAGAAAGAAACGCTCCAGGCCCAAGCAGGAAAGGGAGCT 600  
Qy 668 TCACCTGTTGAGTTTGCATATAGATCTTAATGAACCTACATCTGCTTATGCAACCAAGTG 727  
Db 601 TCACCTGTTGAGTTTGCATATAGATCTTAATGAACCTACATCTGCTTATGCAACCAAGTG 660  
Qy 728 TCTTATGGGAGATGATAGGATGTGCAATGAACAGTGTCCAATTTGAATGGTTTCTACTTT 787





Db 348 GCGCAGCAGGAGCTGGCGGACACACAGTGGGCAACAGCGGCAAGGTTGGCGGACAGGCG 407  
Qy 477 AGCAAGATGGATTCCAGCCACCAAGATCTTCAGAGAGACCCCGCAG--CGACGG 534  
Db 408 CAATGGCGATGCGGTAGCGAGTCTGACAAAGCCCAACAGCAAGCGCTCAGCGGCGGACGG 467  
Qy 535 GACCAAGTGAAGCGCGTGAATTTATGTCAATGGGCAATGGGATTTGAAGACTGTGATGATCA 594  
Db 468 CAACAAGCAACCGTGAGNACCGGTCAGCAACACAGCCAGCAGCGGCGCTCGGG 527  
Qy 595 GCCACCTAAAGAAAAGAAATCCAAAGTCAGCAAAAGAAAAGAAACGCTCCAAAGCCCAAGCA 654  
Db 528 CACACCAAGGAAGAGGCGCAAGACCTCCAAGAAAGAAAGAGCGCTCCAAAGCCCAAGGC 587  
Qy 655 GGAAGGGAAGCTTCACCTGTTCAGTTTGGATTTGCAATAGATCCTAATGAACCTACACTGCTT 714  
Db 588 GGAGCGAGAGGCGTCCCTGCGGACCTCCCATCGACCCCAACGACCGTACTGTCT 647  
Qy 715 ATGCAACCAAGTCTCTTATGGGAGATGATAGGATGTGCAATGAACAGTGTCCAATTGA 774  
Db 648 GTGCAACAGGCTCTCTATGGGAGATGATCGGCTGGCAACAGCAGTGTGCCCATCGA 707  
Qy 775 ATGGTTTCATTTTCATGTGTTTCACTTACCTATAAACCAAGGGGAAATGGTATTGCC 834  
Db 708 GTGGTTTCACTTCTCTGCGTGGGCTCAATATAAACCAAGGGCAAGTGTGTCTGCC 767  
Qy 835 AAGTGCAGGGAGATAATGAGAAAACATGGAACAAAGTACTGAAAGACAAAAGGA 894  
Db 768 CAAAGTCCGGGGGGAACAGAGAACCAATGGAACAAAGCCCTGGAGAAATCCAAAAGAA 827  
Qy 895 TAG 897  
Db 828 GAG 830

## RESULT 5

US-09-764-877-3454  
; Sequence 3454, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3454  
; LENGTH: 8487  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-3454

Query Match 17.0%; Score 183.4; DB 10; Length 8487;  
Best Local Similarity 57.0%; Pred. No. 5.5e-33;  
Matches 395; Conservative 0; Mismatches 266; Indels 32; Gaps 2;

Qy 236 CAAGAAACGTTAAAGGAAATGTATGTCTACGAAAAATATAGAAAGAGATGATTTA 295  
Db 6229 CCAGAGATCTTGAAGGAGCTAGAGATGTCTACGAGCGCTTCACTCGGAGACAGCGG 6288  
Qy 296 AACCAAGAAACGCTTACAGAGCTTCTCCAGAGAGCTAATTAATAGTCAAGAAATTG 355  
Db 6289 GCGCAGAAAGCGGCGGATGTGCACTGTGTGACGCGCGCTGATCCGACGCCAGGAGCTG 6348  
Qy 356 GGAAGTGAATAATACAGATTTGTACACAATCTCGAATTTGTTGGAATCGGCAAGA 415  
Db 6349 GCGCAGAGAAATCAGATCTGAGCCAGATGGTGGAGTGTGGAGAACCCGACGCGG 6408  
Qy 416 CAAATGGAGTTACACTCACAGTGTCTTCCAAAGATC----- 449  
Db 6409 CAGGTGACAGCCAGCTGGAGCTGTTTCGAGGCGCAGCAGGAGCTGGCGACACAGCGGG 6468

Qy 450 -----CTGCTGAAGTGAACGAGCCTCAGATAAAGCAAGATGGATTCCAGCGCAACCA 504  
Db 6469 AACACGGCAGAGCTGGCGGACAGCGCCAAAGCGGAGCGCAGCGAGCTGACACAG 6528  
Qy 505 AAGATCTTTCAAGAAACCCCGCAGCGGACCGAGTGAAGCCGTTATGTACAT 564  
Db 6529 CCCAACAGCAAG-CGCTCAGCGCGCAGCGCAACACGAGAACCGTGAGAACGCGTCCAG 6587  
Qy 565 GCAAAATGGGATTGAAGACTGTGATGATCAGCCACTTAAGAAAAGAAATCCAAAGTCAGC 624  
Db 6588 CAACACACGACCAACGACGCGCGCTCGGGCAACCCCAAGGAGAAAGGCCCAAGACCTC 6647  
Qy 625 AAGAAAAAGAAACCTCCAAAGGCAAGCAGGAAAGGAAAGCTTCACCTGTTGAGTTTGC 684  
Db 6648 CAAGAAGAAAGACCGCTCAAGGCCAAGCGGAGGAGAGGCGTCCCTATGCGGAGATGAT 6707  
Qy 685 AATAGATCTTATGAACCTACATACCTATGCTTATGCAACCAAGTGTCTTATGGGAGATGAT 744  
Db 6708 CATCGACCCCAAGAACCCACGTAAGTGTGTGCAACCAAGGCTCTCTATGCGGAGATGAT 6767  
Qy 745 AGGATGTGACAAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCACTGTTTCACTTAC 804  
Db 6768 CGGCTGCGACACGACGAGTGCCTCATCGAGTGGTTCCACTTCTCGTGGGCGCTCAA 6827  
Qy 805 CTATAAACCAAGGGGAAATGGTATTGCCAAAGTGCAGGGGAGAGATAATAGAAAAACAAT 864  
Db 6828 TCATAAACCAAGGGCAAGTGTGCTGCTCCAAAGTGCAGGGGAGAGAACGAGAACCAT 5887  
Qy 865 GCACAAAAGTACTGAAAAACAAAAAGGATAG 897  
Db 6888 GCACAAAGCCCTGGAGAAATCCAAAAAGAGAG 6920

## RESULT 6

US-09-968-653A-2  
; Sequence 2, Application US/09968653A  
; Publication No. US20030073084A1  
; GENERAL INFORMATION:  
; APPLICANT: Guckov, Andrey V  
; Garkavstev, Igor  
; Riabowol, Karl  
; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/968,653A  
; FILING DATE: 01-Oct-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/006,783A  
; FILING DATE: 15-JAN-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. US20030073084A1Alan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 97,837  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:



```
RESULT 8
US-10-117-722-435
; Sequence 435, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28CIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 435
; LENGTH: 1864
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (164)..(1420)
US-10-117-722-435

Query Match      8.5%; Score 91.6; DB 13; Length 1864;
Best Local Similarity 64.8%; Pred. No. 2.1e-11;
Matches 136; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 674 GTTGAGTTTGCATAGATCCTTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTAT 733
DB 1214 GTTGATTTGGACTTTACGACCCCAATGAACCTCGATCTGCAATTTGTAATCAGGTATCTTAT 1273

QY 734 GGGGAGATGATAGATGTGACATGAACAGTGTCCCAATGAATGGTTTCACITTTTCATGT 793
DB 1274 GGTGAGATGGTGGATGTATACCAAGATGGCTTATAGAAATGGTTCCATTTATGCTGTC 1333

QY 794 GTTTCACCTTACCTATTAACCAAGGGGAAATGGTATGCCCCAAAGTCGAGGGGAGATAAT 853
DB 1334 GTTGGATTGACAGAGGCGACCAAAAGGCAAAATGGTACTGTCCACAGTGCAGTGTGCAATG 1393

QY 854 GAGAAACAAATGACAAAGTACTGAAAAG 883
DB 1394 AAGAGAAGAGCGCAGCACACAAATAAAGG 1423

RESULT 9
US-10-037-270-435
; Sequence 435, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
```

```
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 435
; LENGTH: 1864
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (164)..(1420)
US-10-037-270-435

Query Match      8.5%; Score 91.6; DB 15; Length 1864;
Best Local Similarity 64.8%; Pred. No. 2.1e-11;
Matches 136; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 674 GTTGAGTTTGCATAGATCCTTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTAT 733
DB 1214 GTTGATTTGGACTTTACGACCCCAATGAACCTCGATCTGCAATTTGTAATCAGGTATCTTAT 1273

QY 734 GGGGAGATGATAGATGTGACATGAACAGTGTCCCAATGAATGGTTTCACITTTTCATGT 793
DB 1274 GGTGAGATGGTGGATGTATACCAAGATGGCTTATAGAAATGGTTCCATTTATGCTGTC 1333

QY 794 GTTTCACCTTACCTATTAACCAAGGGGAAATGGTATGCCCCAAAGTCGAGGGGAGATAAT 853
DB 1334 GTTGGATTGACAGAGGCGACCAAAAGGCAAAATGGTACTGTCCACAGTGCAGTGTGCAATG 1393

QY 854 GAGAAACAAATGACAAAGTACTGAAAAG 883
DB 1394 AAGAGAAGAGCGCAGCACACAAATAAAGG 1423

RESULT 10
US-09-738-973-146
; Sequence 146, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 146
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-146
```

Query Match 7.7%; Score 83.6; DB 10; Length 451;  
Best Local Similarity 58.6%; Pred. No. 7.3e-10;  
Matches 163; Conservative 0; Mismatches 114; Indels 1; Gaps 1;  
QY 169 CGTGGAGTCCTGCCCCACACATGACAGAGAACTGTCTGTCTGCGGAGCTGGACAA 228  
Db 16 CATCGAGTCCTCCCTTTTCGACTTGCAGAGAAATGTCTCGCTGATCGGGAGATCGACGC 75  
QY 229 CAAATATCAAGAAACGTTAAAGGAAATGTATGTCTACGAAATATTAAGAAAGA 288  
Db 76 GAATACCAAGAGATCTTGAGAGAGCTAGACGAGTGTACAGCGCTTCAGTCCGAGAC 135  
QY 289 TGATTTAAACCAAGAAACGTTACAGAGCTTCTCCAGAGACGCACTAATTAATAGTCA 348  
Db 136 AGACGGGGCGCAGAGCGGGGATGTCTGCTGTGTGACGCGCGCTGATCCGCA-CCA 194  
QY 349 AGAATTTGGAGATGAAAAATACAGATTTTACAAATGCTCGAATTTGGTGAATAATCG 408  
Db 195 GGAGCTGGGCGACGAGAGATCCAGATCGTGAGCCAGATGGTGGAGCTGGTGGAGAACCG 254  
QY 409 GCGAAGCAAAATGGAGTTACACTCACAGTGTTCCTCAAG 446  
Db 255 CACGGCGAGGTGGACAGCCACGTGGAGCTGTTTCGAGG 292

RESULT 11  
US-09-854-133-146  
; Sequence 146, Application US/09854133  
; Publication No. US20020183499A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C10  
; CURRENT APPLICATION NUMBER: US/09/854,133  
; CURRENT FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 735  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 146  
; LENGTH: 451  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-854-133-146

Query Match 7.7%; Score 83.6; DB 10; Length 451;  
Best Local Similarity 58.6%; Pred. No. 7.3e-10;  
Matches 163; Conservative 0; Mismatches 114; Indels 1; Gaps 1;  
QY 169 CGTGGAGTCCTGCCCCACACATGACAGAGAACTGTCTGTCTGCGGAGCTGGACAA 228  
Db 16 CATCGAGTCCTCCCTTTTCGACTTGCAGAGAAATGTCTCGCTGATCGGGAGATCGACGC 75  
QY 229 CAAATATCAAGAAACGTTAAAGGAAATGTATGTCTACGAAATATTAAGAAAGA 288  
Db 76 GAATACCAAGAGATCTTGAGAGAGCTAGACGAGTGTACGAGCGCTTCAGTCCGAGAC 135  
QY 289 TGATTTAAACCAAGAAACGTTACAGAGCTTCTCCAGAGACGCACTAATTAATAGTCA 348  
Db 136 AGACGGGGCGCAGAGCGGGGATGTCTGCTGTGTGACGCGCGCTGATCCGCA-CCA 194  
QY 349 AGAATTTGGAGATGAAAAATACAGATTTTACAAATGCTCGAATTTGGTGAATAATCG 408  
Db 195 GGAGCTGGGCGACGAGAGATCCAGATCGTGAGCCAGATGGTGGAGCTGGTGGAGAACCG 254  
QY 409 GCGAAGCAAAATGGAGTTACACTCACAGTGTTCCTCAAG 446  
Db 255 CACGGCGAGGTGGACAGCCACGTGGAGCTGTTTCGAGG 292

RESULT 12  
US-10-144-649A-146  
; Sequence 146, Application US/10144649A  
; Publication No. US20030118599A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Ligu  
; APPLICANT: Algate, Paul A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C11  
; CURRENT APPLICATION NUMBER: US/10/144,649A  
; CURRENT FILING DATE: 2002-08-21  
; NUMBER OF SEQ ID NOS: 749  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 146  
; LENGTH: 451  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-144-649A-146

Query Match 7.7%; Score 83.6; DB 15; Length 451;  
Best Local Similarity 58.6%; Pred. No. 7.3e-10;  
Matches 163; Conservative 0; Mismatches 114; Indels 1; Gaps 1;  
QY 169 CGTGGAGTCCTGCCCCACACATGACAGAGAACTGTCTGTCTGCGGAGCTGGACAA 228  
Db 16 CATCGAGTCCTCCCTTTTCGACTTGCAGAGAAATGTCTCGCTGATCGGGAGATCGACGC 75  
QY 229 CAAATATCAAGAAACGTTAAAGGAAATGTATGTCTACGAAATATTAAGAAAGA 288  
Db 76 GAATACCAAGAGATCTTGAGAGAGCTAGACGAGTGTACGAGCGCTTCAGTCCGAGAC 135  
QY 289 TGATTTAAACCAAGAAACGTTACAGAGCTTCTCCAGAGACGCACTAATTAATAGTCA 348  
Db 136 AGACGGGGCGCAGAGCGGGGATGTCTGCTGTGTGACGCGCGCTGATCCGCA-CCA 194  
QY 349 AGAATTTGGAGATGAAAAATACAGATTTTACAAATGCTCGAATTTGGTGAATAATCG 408  
Db 195 GGAGCTGGGCGACGAGAGATCCAGATCGTGAGCCAGATGGTGGAGCTGGTGGAGAACCG 254  
QY 409 GCGAAGCAAAATGGAGTTACACTCACAGTGTTCCTCAAG 446  
Db 255 CACGGCGAGGTGGACAGCCACGTGGAGCTGTTTCGAGG 292

RESULT 13  
US-09-801-368-257  
; Sequence 257, Application US/09801368  
; Patent No. US20020128250A1  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Cali, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: No. US20020128250A1man, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/801,368  
; CURRENT FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 16:38:42 ; Search time 93 Seconds

(without alignments)

5125.740 Million cell updates/sec

Title: US-09-513-365A-2

Perfect score: 1080

Sequence: 1 gggccgcggcggtgcatg.....aaaaaaaaaaaaaaaaaa 1080

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*
  - 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*
  - 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*
  - 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*
  - 5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*
  - 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1046	96.9	1078	4	US-09-601-478-7
2	838.4	77.6	840	4	US-09-601-478-6
3	825.2	76.4	1154	4	US-09-484-970B-81
4	237.4	22.0	873	3	US-09-006-783A-4
5	183.8	17.0	1902	2	US-09-258-257-1
6	183.8	17.0	1902	2	US-09-258-371-1
7	183.8	17.0	1902	3	US-08-569-721A-1
8	183.8	17.0	1902	3	US-08-751-230-1
9	183.8	17.0	1902	3	US-09-499-082-1
10	183.8	17.0	1902	3	US-09-258-372-1
11	183.8	17.0	2061	2	US-09-258-371-9
12	183.8	17.0	2061	3	US-08-751-230-9
13	183.8	17.0	2061	3	US-09-499-082-9
14	183.8	17.0	2061	3	US-09-258-372-9
15	182.2	16.9	2061	3	US-09-006-783A-2
16	182.2	16.9	2061	4	US-09-159-871-1
17	174.4	16.1	633	3	US-09-006-783A-6
18	95.6	8.9	678	3	US-09-195-286-3
19	95.6	8.9	699	3	US-09-195-286-2
20	91.6	8.5	1864	4	US-09-620-312D-435
21	83.6	7.7	451	4	US-09-370-838-146
22	72.4	6.7	7218	1	US-08-232-463-14
23	50.6	4.7	860	1	US-07-847-010-18
24	49	4.5	593	3	US-09-385-982-262
25	46.8	4.3	6124	4	US-08-213-419B-3
26	46.2	4.3	1189	1	US-07-781-034-4
27	46.2	4.3	1189	5	PCT-US92-08328-4

28	45.8	4.2	19124	2	US-08-487-826B-13	Sequence 13, Appl
29	43.8	4.2	1664976	4	US-08-916-421B-1	Sequence 1, Appli
30	43.8	4.1	2327	4	US-09-852-067-1	Sequence 1, Appli
31	43.8	4.1	3489	2	US-08-728-323A-1	Sequence 1, Appli
32	43.8	4.1	3489	4	US-09-298-568-1	Sequence 1, Appli
33	43.8	4.1	3489	4	US-09-410-369-1	Sequence 1, Appli
34	43.8	4.1	32207	3	US-08-770-379-20	Sequence 20, Appl
35	43.8	4.1	32207	3	US-08-757-669A-20	Sequence 20, Appl
36	43.8	4.1	32207	4	US-09-230-371A-20	Sequence 20, Appl
37	43.4	4.0	1147	1	US-08-665-716-1	Sequence 1, Appli
38	43.2	4.0	3825	3	US-09-208-742-3	Sequence 3, Appli
39	43	4.0	1286	4	US-09-489-847-124	Sequence 124, App
40	43	4.0	1342	4	US-09-489-847-89	Sequence 89, Appli
41	43	4.0	3483	4	US-09-130-491-3	Sequence 3, Appli
42	42.8	4.0	603	3	US-09-385-982-251	Sequence 251, App
43	42.8	4.0	774	3	US-08-956-307B-12	Sequence 12, Appl
44	42.8	4.0	778	3	US-08-956-307B-11	Sequence 11, Appl
45	42.8	4.0	3377	6	US-08-345-16	Patent No: 5198345

ALIGNMENTS

RESULT 1  
US-09-601-478-7  
; Sequence 7, Application US/09601478  
; Patent No. 6403785  
; GENERAL INFORMATION:  
; APPLICANT: Oetuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human INGLL gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/09/601,478  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1078  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
; NAME/KEY: CDS  
; LOCATION: (92)..(931)  
US-09-601-478-7

Query Match	96.9%	Score	1046	DB	4	Length	1078
Best Local Similarity	99.5%	Pred. No.	4.4e-264				
Matches	1049	Conservative	0	Mismatches	5	Indels	0
Gaps	0						
QY	1	GGGCGCGCGCGGTCATGTCGGCTGCGGTCGATGCGGAGCGCGCGCGCGCGGAT	60				
Db	25	GGGCGCGCGCGGTCATGTCGGCTGCGGTCGATGCGGAGCGCGCGCGCGGAT	84				
QY	61	CGGCGAGATGTTAGGCGAGCAGCAGCACTGTACTCGTCGGCTGCGCTCTGACCGG	120				
Db	85	CGGCGAGATGTTAGGCGAGCAGCAGCACTGTACTCGTCGGCTGCGCTCTGACCGG	144				
QY	121	GGAGCGAGCGCGCTGCTCACCTGCTACCTGTCAGGAGTACCTTTAGTGGTGGAGT	180				
Db	145	GGAGCGAGCGCGCTGCTCACCTGCTACCTGTCAGGAGTACCTTTAGTGGTGGAGT	204				
QY	181	GGCCACGACATGCGAGGAACTGTCTGTGTCGAGAGCTGGACCAAAATATCAAGA	240				
Db	205	GGCCACGACATGCGAGGAACTGTCTGTGTCGAGAGCTGGACCAAAATATCAAGA	264				
QY	241	AACGTTAAAGGAATGATGTCCTACGAAAAATATAAGAAAGAGATGATTTAAACCA	300				







168 CTACGAGCGCTTCAGTCGCGAGACACGCGGGCGCAGAGCGCGGATGTCGACCTGTGT 227  
325 CCGAGAGACCTAATTAATAGTCAAGAAATGGAGATGAAATAACAGATTGTACACA 384  
228 GCAGCGCGCGCTGATCCGAGCGCAGGAGCTGGGCGAGAGATCCAGATCGTGAGCCA 287  
385 AATGCTCGAATGGTGGAAATCGGCGACAGCAAAATGGAGTTACATCAAGTGTTCCTCA 444  
288 GATGGTGGAGCTGGTGAGAACCGCAGCGGCGAGGTGGACCGCACCGTGAGGTGTTCGA 347  
445 AGATC-----CTGCTGAAAGTGAACGAGCCTCAGATA 476  
348 GCGCAGCAGGAGCTGGCGACACAGTGGGCAACAGCGGCAAGTTGGCGCGACAGGCC 407  
477 AGCAAAAGATGATTCAGCCAAACAGAAAGATCTTCAAGAGAACCCCGGAG--GAGCG 534  
408 CAATGGCGATGCGGTAGCGCAGTCTGACAAAGCCCAACAGCAAGCGCTCACGGCGGAGCG 467  
535 GACCACTGAAGCGCGTATTTATGTACATGGCAATGGATTTGAAGACTGTGATGATCA 594  
468 CACAAACGAGAACCGTGAGAACCGGCTCCAGAACCCAGCACGACGCGCGCTCGGG 527  
595 GCCACCTAAAGAAAGAAATCCAAAGTCAGCAAAAGAAAGAAAGAAAGAAAGAAAGAAAG 654  
528 CACACCCAGGAGAGAGGCGCAAGACCTCCAGAGAGAGAGAGCGCTCCAGGCGCAAGGC 587  
655 GGAAGGAGAGCTTACCTGTTGAGTTGCAATAGATCTTAATGAACCTACATCTGCTT 714  
588 GGAGCGAGAGCGGCTCCCTGCGGACCTCCCAATCGACCCCAACCAACCCACGCTACTGTCT 647  
715 ATGCAACCAAGTCTTTATGGGAGATGATAGTGTGACATGAACAGTGTCCAATTGA 774  
648 GTGCAACCAAGTCTCTTATGGGAGATGATCGGCTCGGACACGAGTGTGCTTCTGCT 707  
775 ATGCTTTCATCTTTCATGTTTCACTTACTATAAACCAGGAGAAATGGTATTGCCCC 834  
708 GTGCTTCCATCTCTGTCGCTGGGCTCATCATATAACCCAGGCGAAGTGTGCTGCTC 767  
835 AAGTGCAGGGAGATAATGAGAAAACAAATGGACAAAGTACTGAAAGACAAAAGGA 894  
768 CAAGTGCAGGGGAGAACGAGAAAGACCATGGACAAAGCCCTGGAGAAATCCCAAAAAGA 827  
895 TAG 897  
828 GAG 830

RESULT 5

US-09-258-257-1  
Sequence 1, Application US/09258257  
Patent No. 5965398  
GENERAL INFORMATION:  
APPLICANT: GARKAVTSEV, Igor  
APPLICANT: RIABOWOL, Karl  
TITLE OF INVENTION: DNA SEQUENCE ENCODING A TUMOR  
TITLE OF INVENTION: SUPPRESSOR GENE  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/258,257  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/569,721  
FILING DATE: 08-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-128  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 854-7400  
TELEFAX: (650) 854-8275  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1902 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 109..738  
US-09-258-257-1  
Query Match 17.0%; Score 183.8; DB 2; Length 1902;  
Best Local Similarity 57.0%; Pred. No. 2.2e-38;  
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;  
QY 237 AAGAAACGTTAAAGAAATTTGATGTCTACGAAATATATAGAAAGAAAGATGATTTAA 296  
DB 35 AACAGATCTCTGAAGGAGCTAGACGAGTGTCTACGAGCGCTTCAGTCGCGAGACAGCGGG 94  
QY 297 ACCAGAAGAAACGCTACAGCAGCTTCTCCAGAGAGCCTAATTAATAGTCAAGAAATTTG 356  
DB 95 CCGAGAAGCGCGGATGCTGCACTGTGTGCGCGCGCTGATCCGACCGAGGAGCTGG 154  
QY 357 GAGATGAAATATACAGATTTGTTACAAATCTCGAATTTGGTGGAAATTCGGGCAAGAC 416  
DB 155 GCGAGAGAGATCCAGATCGTGAGCCAGATGTTGAGGAGTGGTGGAGAACCGCACCGGC 214  
QY 417 AAATGGAGTTACCTCACTCAGTGTCTTCCAAAGATC-----449  
DB 215 AGGTGACAGCCACGCTGGAGCTGTTTCGAGGCGCAGCAGGAGCTGGGCGCACACAGTGGCA 274  
QY 450 -CTGCTGAAGTGAACGAGCCTCAGATAAGCAAGATGATTTCCAGCCAAACAGAAAGA 508  
DB 275 ACAGCGGCAAGTTGGCGGACAGGCCCAATGGCGATCGGTAGCGCAGTCTGACAAGC 334  
QY 509 TCTTCAAGAA--GACCCCGCAGCGAGCGGACCGAGTGAAGCGGTGATTTATGTCAATGG 566  
DB 335 CCAACAGCAAGCGCTCACGGCGCAGCGCAACCAAGAGAACCGTGAGAACCGGTCCAGCA 394  
QY 567 CAAATGGGATTAAGAACTGTGATGATCAGCCACTTAAGAAAGAAATCCAGTCAAGCAA 626  
DB 395 ACCACGACCAACGACGCGCTCGGGCACACCCCAAGGAGAAAGAGGCGCAAGACCTCCA 454  
QY 627 AGAAAAGAAACGCTCCAGGCGCAAGCAGAAAGGAAAGCTTCACTGTTGAGTTGCA 686  
DB 455 AGAAGAAAGAAAGCGCTCCAAAGGCGGAGGCGGAGGCGTCCCTGCCGACCTCCCCA 514  
QY 687 TAGATCCTAATGAACCTACATACTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746  
DB 515 TCGACCCCAACGACCCAGTACTGCTGTGCAACACAGGTCTCTATGGGAGATGATCG 574  
QY 747 GATGTGACAAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCACTGTTTCACTTACCT 806  
DB 575 GCTGCGACAAACGACGAGTGCCTCCATCGAGTGGTTCACCTTCTCGTGGGCGTCAATC 634  
QY 807 ATAAACCAAGGGGAAATGGTATTGCCCAAGTCAGGGGAGATATAGAAACCAATGG 866  
DB 635 ATAAACCAAGGGGAAATGGTATTGCCCAAGTCAGGGGAGATATAGAAACCAATGG 694  
QY 867 ACAAAAGTACTGAAAAAGACAAAAGGATAG 897





Db 575 GCTGCGACACGACGAGTGTCCCATCAGTGGTTCCACTTCTCGTGGGGCTCAATC 634  
QY 807 ATAAACCAAGGGAATGTAATTTGCCAAAGTGCAGGGGAGAGATAATAGAAAACAAATGG 866  
Db 635 ATAAACCAAGGGAATGTAATTTGCCAAAGTGCAGGGGAGAGATAATAGAAAACAAATGG 694  
QY 867 ACAAAGTACTGAAAGACAAAAGGATAG 897  
Db 695 ACAAGCCCTGGAGAAATCCAAAAGAGAG 725

## RESULT 9

US-09-499-082-1

; Sequence 1, Application US/09499082

; Patent No. 6143522

; GENERAL INFORMATION:

; APPLICANT: Helbing, Caren C.

; APPLICANT: Riabowol, Karl

; APPLICANT: Johnston, Randall N.

; APPLICANT: Garkavtsev, Igor

; TITLE OF INVENTION: METHODS OF MODULATING APOPTOSIS

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Burns, Doane, Swecker &amp; Mathis

; STREET: 699 Prince Street

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-1404

; COMPUTER READABLE FORM: disk

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; FILING DATE: US/09/499,082

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/828,158

; FILING DATE: 27-MAR-1997

; APPLICATION NUMBER: US 08/751230

; FILING DATE: 15-NOV-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/569721

; FILING DATE: 08-DEC-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Mool, Leslie A.

; REGISTRATION NUMBER: 37,047

; REFERENCE/DOCKET NUMBER: 028722-148

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-854-7400

; TELEFAX: 650-854-8275

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1902 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 109..741

US-09-499-082-1

Query Match 17.0%; Score 183.8; DB 3; Length 1902;  
Best Local Similarity 57.0%; Pred. No. 2.2e-38;  
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;  
QY 237 AGAAGCGTTAAAGGAATTTGATGCTTACGAAAATATAAGAAAGATGATTTAA 296  
Db 35 AACAGATCTCTGAAGGAGTAGACGAGTGCTTACGAGCGCTTCAGTCGCGAGACAGACGGG 94

QY 297 ACCAGAGAAAACGTTCTACAGCAGCTTCTCCAGAGACACTAATAATAGTCAAGAAATTGG 356  
Db 95 CGCAGAACGGGGGAGTGTCTGCACTGTGTGTCAGCGCGCTGATCCGAGCCAGGAGCTGG 154  
QY 357 GAGATGAAAAATACAGATTGTTACACAATGCTCGAATTTGGTGGAAATTCGGGCAAGAC 416  
Db 155 GCGACGAGAGATCCAGATCGTGAGCCAGATGGTGGAGCTGTGTGGAGAACCGCACCGCGC 214  
QY 417 AAATGGAGTTTACACTCACAGTGTTCCTCAAGATC----- 449  
Db 215 AGTGGACAGCCACGTTGGAGCTGTTGAGCGCGCAGGAGCTGGGCGACACAGTGGGCA 274  
QY 450 -CTGCTGAAAGTGAACGAGCCTCAGATAAAGAAAGATGATTCAGCCCAACAGAAAAGA 508  
Db 275 ACAGCGGCAAGGTTGGCGCGGACAGGCCCAATGGCGTAGCGGAGCTCTGACAAAGC 334  
QY 509 TCTTCAAGAA--GACCCCGCAGCAGCGGACCAAGTGAAGCCGTGATTTATGTCACATGG 566  
Db 335 CCAACAGCAAGCGCTCACGGCGGACGCGCAACAGAGAACCGTGAGAACGCGTCCAGCA 394  
QY 567 CAAATGGGATTGAAGACTGTGATGATCAGCCACTAAAGAAAAGAAATCCAAAGTCAGCAA 626  
Db 395 ACCACGACACGACGACGGCGCTCGGGCACACCCCAAGAGAGAGAGCCCAAGCCTCCA 454  
QY 627 AGAAAAAGAAACGCTCCAAAGGCCAAGCAGGAAAGGAAAGCTTCACCTGTTGAGTTGCAA 686  
Db 455 AGAAGAAAGAGCGCTCCAAAGGCCAAGCGGAGGAGGCGTCCCTGCGGACCTCCCCA 514  
QY 687 TAGATCCTAATGAACCTACATCTGTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746  
Db 515 TCGACCCCAACGAAACCCAGTACTGTCTGTGCAACCAAGGTCTCTATGGGAGATGATCG 574  
QY 747 GATGTGACAAATGAACAGTGTCCAATTTGAATTTTCACTTTTTCATGTTTCACTTACCT 806  
Db 575 GCTGCGACAAACGACGAGTGCCCAAGTGTGAGTGTTCACCTTCTCTGCTGGGGGCTCAATC 634  
QY 807 ATAAACCAAGGGGAAATGTTATTTGCCAAAGTGCAGGGGAGAGATAATGAGAAAAACAATGG 866  
Db 635 ATAAACCAAGGGGAAATGTTATTTGCCAAAGTGTGTTCCCAAGTGTGCGGGGAGAACGAGAACCATGG 694  
QY 867 ACAAAGTACTGAAAGACAAAAGAGATAG 897  
Db 695 ACAAGCCCTGGAGAAATCCAAAAGAGAG 725

RESULT 10  
US-09-258-372-1  
; Sequence 1, Application US/09258372  
; Patent No. 6238918  
; GENERAL INFORMATION:  
; APPLICANT: Garkavtsev, Igor  
; APPLICANT: Riabowol, Karl  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
; TITLE OF INVENTION: SUPPRESSOR GENE INGI  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/258,372  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/751,230

FILING DATE: 17.0%; Score 183.8; DB 3; Length 1902;  
Best Local Similarity 57.0%; Pred. No. 2.2e-38;  
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;  
Query Match

ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-144  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-7400  
TELEFAX: 415-854-8275  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1902 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
NAME/KEY: CDS  
LOCATION: 109..741  
US-09-258-372-1

Query Match 17.0%; Score 183.8; DB 3; Length 1902;  
Best Local Similarity 57.0%; Pred. No. 2.2e-38;  
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;  
Query Match

237 AACAAACGTTAAAGGAAATTTGATGTTCTACGAAATATATAAGAAAGAGATGATTTAA 296  
|||  
35 AACAGATCTCTGAAGGAGTAGACGAGTGTACGAGCGCTTCAGTCGCGAGACAGACGGGG 94  
|||  
297 ACCAGAGAAACGCTTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGAAATGG 356  
|||  
95 CGCAGAAGCGGGGATGCTGCACTGTGTGCGAGCGCGCTGATCCGAGCAGGAGCTGG 154  
|||  
357 GAGATGAAATAATACAGATGTTTACAAATGCTCGAATGTTGGTGGAAATTCGGGCAAGAC 416  
|||  
155 GCACGAGAAGATCCAGATGCTGAGCGAGCAGATGTTGGAGTGTGTGGAACCGCACCGCG 214  
|||  
417 AATGAGGTTTACACTCACAGTGTTCAGATC----- 449  
|||  
215 AGTGGACACCCACGCTGGAGCTGTTTCAGCGCAGCAGGAGCTGGCGGACACAGTGGGCA 274  
|||  
450 -CTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATGATTCAGCCCAACAGAAAGA 508  
|||  
275 ACAGCGGCAAGTTGGCGGACAGGCCCAATGGCGATGCGTAGCGAGCTGTGCAAGC 334  
|||  
509 TCTTCAAGAA--GACCCCGCAGCAGCGGACAGTGAACCGTGTATTTATGTCATGG 566  
|||  
335 CCAACAGCAGCGCTCACGCGCGCAGCGCAACACAGAGAACCGTGAGAACCGCTCCAGCA 394  
|||  
567 CAATGGGATTGAAGACTGTGATGATCAGCCACCTTAAAGAAAGAAATCCAGCTCAGCAA 626  
|||  
395 ACCACAGCACGACGAGCGGCGCTCGGGCACCCCAAGGAGAGAGGCGCAAGCCTCCA 454  
|||  
627 AGAAAAGAAACGCTCCAGGCGCAAGCAGGAAAGGAAAGCTTCACCTGTTGAGTTTGC 686  
|||  
455 AGAAGAAGAGCGCTCCAGGCGCAAGCGGAGGAGGAGGCTCCCTGCGGACCTCCCA 514  
|||  
687 TAGATCCTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746  
|||  
515 TCACGCCCAACGAAACCGTACTGCTGTGCAACCAAGGCTCTCTATGGGAGATGATCG 574  
|||  
747 GATGTGACATGAACAGTGTCCAAATGAATGTTTTCATTTTCACTGTTTTCATTTACCT 806  
|||  
575 GCTGCGACACGACGAGTGTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 834  
|||  
807 ATAAACCAAGGGAATGTTATGCGCAAGTGTGCGGAGGAGATTAATGAGAAACCAATGG 866  
|||  
635 ATAAACCAAGGCGAGTGTGCTGCTCCCAAGTGTGCGGAGGAGACGAGAACCATGG 694  
|||  
867 ACAAAGTACTGAAAGACAAAGGATGAG 897  
|||  
695 ACAAGCCCTGGAGAAATCCAAAAAGAGAG 725  
|||

RESULT 11  
US-09-258-371-9  
Sequence 9, Application US/09258371  
Patent No. 5986078  
GENERAL INFORMATION:  
APPLICANT: Garkavtsev, Igor  
APPLICANT: Riabowol, Karl  
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
TITLE OF INVENTION: SUPPRESSOR GENE INGI  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/258.371  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/751,230  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-144  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-7400  
TELEFAX: 415-854-8275  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16..900  
US-09-258-371-9

Query Match 17.0%; Score 183.8; DB 2; Length 2061;  
Best Local Similarity 57.0%; Pred. No. 2.3e-38;  
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;  
Query Match

237 AAGAAACGTTAAAGGAAATTTGATGTTCTACGAAATATATAAGAAAGAGATGATTTAA 296  
|||  
194 AACAGATCTCTGAAGGAGCTAGACGAGTGTACGAGCGCTTCAGTCGCGAGACAGACGGGG 253  
|||  
297 ACCAGAGAAACGCTTACAGCAGCTTCTCCAGAGAGCAGTAAATTAATAGTCAAGATTCG 356  
|||  
254 CGCAGAAGCGCGGATGCTGCACTGTGTGCGAGCGCGCTGATCCGACGCCAGGAGCTGG 313  
|||  
357 GAGATGAAATAATACAGATTTGTTACAAATGCTCGAATTTGTTGAAAAATTCGGGCAAGAC 416  
|||  
314 GCGACGAGAGATCCAGATCTGTGAGCCAGATGTTGGAGCTGTTGGAGAACCGCACCGCG 373  
|||  
417 AATGAGGTTTACACTCACAGTGTTCCTCAAGATC----- 449  
|||  
374 AGTGGACAGCCACGCTGGAGCTGTTTCGAGCGCGCAGCAGGAGCTGGCGGACAGTGGGCA 433  
|||  
450 -CTGCTGAAGTGAACGAGCCTCAGATAAAGCAAGATGATTCAGCCCAACAGAAAGA 508  
|||  
434 ACAGCGGCAAGGTTGGCGCGGACAGGCCCAATGGGCGATCGGTAGCGCAGTCTTGAAGC 493  
|||  
509 TCTTCAAGAA--GACCCCGCAGGCGAGCCAGTGAAGCGGTGATTTATGTGCATG 566  
|||







```

; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/499,082
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,158
; FILING DATE: 27-MAR-1997
; APPLICATION NUMBER: US/08/751230
; FILING DATE: 15-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/569721
; FILING DATE: 08-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mooi, Leslie A.
; REGISTRATION NUMBER: 37,047
; REFERENCE/DOCKET NUMBER: 028722-148
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-7400
; TELEFAX: 415-854-8275
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..900
; US-09-499-082-9

Query Match 17.0%; Score 183.8; DB 3; Length 2061;
Best Local Similarity 57.0%; Pred. No. 2.3e-38;
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AACAAAGCTTAAAGGAAATGATGCTTACGAAATAATAAGAAAGAGATGATTAA 296
DB 194 AACAGATCTGAGGAGCTAGACAGTGTCTACGCGCTTCAGTCGCGACAGACGGGG 253
QY 297 ACCAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCCTAAATTAATAGTCAAGAAATGG 356
DB 254 CGCAGAAGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
QY 357 GAGATGAAATAATACAGATGTTTACAAATGCTGAAATGGTGGAAATTCGGGCAAGAC 416
DB 314 CGCAGCAGAAAGATCCAGATCGTGAAGCAGATGCTGAGAGCTGCTGAGAAACCGCAGCG 373
QY 417 AATGAGGATTACACTACAGTGTTCACAGATC----- 449
DB 374 AGTGGACACCCACGTGGAGCTGTTTCAGCGCAGCAGAGAGCTGGGCGACAGTGGGCA 433
QY 450 -CTGCTGAAAGTGAACGAGCCTCAGATPAAAGCAAGATGGAATTCAGCCCAACAGAAAGA 508
DB 434 ACAGCGCAAGTTGGCGGACAGGCGCAATGGCGATGCGGTAGCGCAGTCTGACAAGC 493
QY 509 TCTTCAAGAA--GACCCCGCAGCGGACCGGACCAAGTGAACCGGTGATTTATGTCATGG 566
DB 494 CCAACAGCAAGCGCTCAACCGCGGACGCGCAACACGAGAACCGGTGAGAACCGGTCCAGCA 553
QY 567 CAATGGGATTGAAGACTGTGATGATCAGCCACTTAAAGAAAGAAATCCAACTCAGCAA 626
DB 554 ACCACGACACGACGACGCGCTTCGGGCAACCCAGAGGAGAAAGGCCCAAGACCTCCA 613
QY 627 AGAAAGAAACGCTCCAAAGCCCAAGCAGGAAAGGGAAGCTTCACTGTTGAGTTTGCAG 686
DB 614 AGAAGAAGAGCGCTCCAAAGGCCCAAGCGCGAGGAGGCGTCCCTGTCGCGACAGACGGGG 673

687 TAGATCCTAATGAACCTACATATCTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746
DB 674 TCGACCCCAACGAAACCCACGTACTCTCTGCAACACGAGTCTCTTATGGGAGATGATCG 733
QY 747 GATGTGACAATGAACAGTGTCCAATGAATGATGTTTCACTTTTTCATGTGTGTTTCACTTACCT 806
DB 734 GCTGCGACAAACGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
QY 807 ATAAACCAAGGGGAAATGTTATCCCAAGTGTGAGGAGATGATGAGAAACCAATGG 866
DB 794 ATAAACCAAGGGGCAAGTGTGTTCTCCCAAGTGTGCGGGGAGAACGAGAACCATGG 853
QY 867 ACAAAGTACTGAAAGAGACAAAGAGGATAG 897
DB 854 ACAAAGCTTGGAGAAATCCAAAGAGAGAG 884

RESULT 14
US-09-258-372-9
; Sequence 9, Application US/09258372
; Patent No. 6238918
; GENERAL INFORMATION:
; APPLICANT: Garkavtsev, Igor
; APPLICANT: Riabowol, Karl
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR
; TITLE OF INVENTION: SUPPRESSOR GENE INGI
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/258,372
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/751,230
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mooi, Leslie A.
; REGISTRATION NUMBER: 37,047
; REFERENCE/DOCKET NUMBER: 028722-144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-7400
; TELEFAX: 415-854-8275
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..900
; US-09-258-372-9

Query Match 17.0%; Score 183.8; DB 3; Length 2061;
Best Local Similarity 57.0%; Pred. No. 2.3e-38;
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AAGAAACGTTAAAGGAAATGATGATGCTTACGAAATAATAAGAAAGAGATGATTAA 296
DB 194 AACAGATCTGAGGAGCTAGACAGTGTCTACGCGCTTCAGTCGCGACAGACGGGG 253
```

QY 297 ACCAGAGAAACGTCTACAGACGCTTCTCAGAGAGCACTAATTANTAGTCAAGATTGG 356  
Db 254 CGCAGAAAGGGGGATGCTGCACTGTGTGAGCGCGCTGATCCGACGAGAGCTGG 313  
QY 357 GAGATGAAAAATACAGATTGTTACACAAATCTCGAATTGGTGGAAATCGGGCAAGAC 416  
Db 314 GCGACGAGAAGATCCAGATCGTGAGCCAGATGGTGGAGCTGGTGGAGAACCGCACCGGC 373  
QY 417 AAATGGAGTTACACTCACAGTGTGTTCCAAAGATC----- 449  
Db 374 AGTGTGACACGACGCTGGAGCTGTTCTGAGCGCAGCAGGAGCTGGCGACACAGTGGCA 433  
QY 450 -CTGCTGAAGTGAACAGCGCTCAGATAAAGCAAGATGGATTCAGCCCAACAGAAAGA 508  
Db 434 ACAGCGGCAAGTGTGGCGGACAGCCCAATGGCGATGGTGGAGCTGTGACAAAGC 493  
QY 509 TCTTCAAGAA--GACCCCGCAGCGAGCGGACAGTGAAGCCGCTGATTTATGTCATCG 566  
Db 494 CCAACAGCAAGCGCTCACGGCGCAGCGCAACACAGAACCGGTGAGAACGGGTCCAGCA 553  
QY 567 CAATGGGATTGAAGACTGTGATGATCAGCCACTAAAGAAAGAAATCCAAAGTCAGCAA 626  
Db 554 ACCACGACCAGCAGCGCGCTCGGCACACCCCAAGGAGAGAGGCCAAGACCTCCA 613  
QY 627 AGAAAAAGAAACGCTCCAAAGGCCAAGCAGAGAAAGGAAAGCTTCACTGTGAGTTGCAA 686  
Db 614 AGAAGAAAGAGCGCTCCAAAGGCCAAGCGGAGGAGCGCTCCCTGCGGACCTCCCA 673  
QY 687 TAGATCCTAATGAACCTACTACTGTTAATGAGTGGTTTTCATCTTTCATCTGTTTACCT 746  
Db 674 TGACCCCAACGACCGATCTGCTGTGCAACAGAGTCTCTATGGGGAGATGATCG 733  
QY 747 GATGTGACAAATGAACAGTGTGTTTCAATGAGTGGTTTTCATCTGTTTCACTTTACCT 806  
Db 734 GCTGCGACACAGCAGAGTGCCCATCGAGTGGTTTTCATCTGCTGCTGGGCTCAATC 793  
QY 807 ATAAACCAAGGGAAATGTTATGTTCCAAAGTGCAGGGGAGATATAGAAAAACAATGG 866  
Db 794 ATAAACCAAGGGCAAGTGTGTTCCAAAGTGCAGGGGAGACGAGAACCATCG 853  
QY 867 ACAAAAGTACTGAAAGACAAAAAGGATAG 897  
Db 854 ACAAGCCCTGGAGAAATCCAAAAAGAGAG 884

## RESULT 15

US-09-006-783A-2  
; Sequence 2, Application US/09006783A  
; Patent No. 6297366  
; GENERAL INFORMATION:  
; APPLICANT: Gudkov, Andrey V  
; APPLICANT: Garkavstev, Igor  
; APPLICANT: Rabinowol, Kari  
; TITLE OF INVENTION: p33/INGI as a Mediator of p53 Signaling  
; TITLE OF INVENTION: Pathway  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/006,783A  
; FILING DATE: 15-JAN-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6297366nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 97,837  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2061 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 16..897  
; US-09-006-783A-2

Query Match 16.9%; Score 182.2; DB 3; Length 2061;  
Best Local Similarity 56.9%; Pred. No. 6.1e-38;  
Matches 393; Conservative 0; Mismatches 268; Indels 30; Gaps 2;  
QY 237 AAGAAACGTTAAGGAATTTGATGATGCTACGAAAAATATAAGAAAGAGATGATTAA 296  
Db 194 AACAGATCTCTGAAGCGCTAGACAGTGTCTACGAGCGCTTCAGTCGCGAGACAGACGGG 253  
QY 297 ACCAGAAGAAACGTCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAAATTGG 356  
Db 254 CGCAGAAAGCGCGGATGCTGCACTGTGTGACGGCGCGCTGATCCGACCGAGAGCTGG 313  
QY 357 GAGATGAAAAATACAGATTGTTACACAAATGCTCGAATTGGTGGAAAAATCGGGCAAGAC 416  
Db 314 GCGACGAGAAGATCCAGATCGTGGAGCCAGATGTTGGAGCTGGTGGAGAACCGCACCGGC 373  
QY 417 AAATGGAGTTACACTCACAGTGTGTTTCCAAAGATC----- 449  
Db 374 AGTGTGACACGACCGCTGGAGCTGTTCTGAGCGCAGCAGGAGCTGGCGCACACAGTGGCA 433  
QY 450 -CTGCTGAAGTGAACAGCGCTCAGATAAAGCAAGATGGATTCAGCCCAACAGAAAGA 508  
Db 434 ACAGCGCAAGTGTGGCGGACAGGCCCAATGGCGATCGGTGAGCGATCTGCAAGC 493  
QY 509 TCTTCAAGAA--GACCCCGCAGCGAGCGGACAGTGAAGCCGCTGATTTATGTCATGG 566  
Db 494 CCAACAGCAAGCGCTCACGGCGCAGCGCAACACAGAGAACCGGTGAGAACCGCTCCAGCA 553  
QY 567 CAATGGGATTGAAGACTGTGATGATCAGCCACTTAAAGAAAGAAATCCAAAGTCAGCAA 626  
Db 554 ACCACGACCAGCAGCGCGCTCGGCACACCCCAAGGAGAGAGGCCAAGACCTCCA 613  
QY 627 AAGAAAGAAACGCTCCAAAGGCCAAGCAGAGAAAGGAAAGCTTCACTGTGAGTTGCAA 686  
Db 614 AGAAGAAAGAGCGCTCCAAAGGCCAAGCGGAGGAGCGCTCCCTGCGGACCTCCCA 673  
QY 687 TAGATCCTAATGAACCTACTACTGTTAATGAGTGGTTTTCATCTTTCATCTGTTTACCT 746  
Db 674 TGACCCCAACGACCGATCTGCTGTGCAACAGAGTCTCTATGGGGAGATGATCG 733  
QY 747 GATGTGACAAATGAACAGTGTGTTTCAATGAGTGGTTTTCATCTTTCATGTTTTCATCTTACCT 806  
Db 734 GCTGCGACACAGCAGAGTGCCCATCGAGTGGTTTTCATCTGCTGCTGGGCTCAATC 793  
QY 807 ATAAACCAAGGGAAATGTTATGTTCCAAAGTGCAGGGGAGATATAGAAAAACAATGG 866  
Db 794 ATAAACCAAGGGCAAGTGTGTTCCAAAGTGCAGGGGAGACGAGAACCATCG 853  
QY 867 ACAAAAGTACTGAAAGACAAAAAGGATAG 897  
Db 854 ACAAGCCCTGGAGAAATCCAAAAAGAGAG 884

Search completed: January 12, 2004, 17:23:53  
Job time : 99 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 16:38:42 ; Search time 2229 Seconds  
(without alignments)  
11776.058 Million cell updates/sec

Title: US-09-513-365A-2  
Perfect score: 1080  
Sequence: 1 ggcggcgcgcgcgtgcatg.....aaaaaaaaaaaaaaaaaaaaa 1080

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_esttc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_est3.\*
- 12: gb\_est4.\*
- 13: gb\_est5.\*
- 14: gb\_est6.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pin.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_mam.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_pbg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gss1.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	828	76.7	2643	11 AK048800	AK048800 Mus muscu
2	827.6	76.6	2812	11 AK083144	AK083144 Mus muscu
3	812.8	75.3	1464	11 AK012716	AK012716 Mus muscu
4	747.8	69.2	912	13 BQ277444	BQ277444 AGENCOURT

C 5	735	68.1	735	12	BM982877	BM982877 UI-CF-ENI
C 6	727.6	67.4	727.6	10	BG184056	BG184056 RST2972 A
C 7	711	65.8	711	10	BE796780	BE796780 601587557
C 8	694.2	64.3	694.2	12	BI548536	BI548536 603191255
C 9	680.8	63.0	680.8	10	BG211544	BG211544 RST31111
C 10	671.8	62.2	671.8	12	BM682789	BM682789 UI-E-EJ1-
C 11	663	61.4	663	12	BM719460	BM719460 UI-E-EJ1-
C 12	636.4	58.9	636.4	12	BI495873	BI495873 df121d04.
C 13	632.2	56.2	632.2	12	BI091047	BI091047 602855235
C 14	607	56.2	607	12	BI460319	BI460319 603201967
C 15	606.4	56.1	606.4	13	BU591073	BU591073 AGENCOURT
C 16	599	55.5	599	9	AV647917	AV647917 AV647917
C 17	563.4	52.2	563.4	9	AM674008	AM674008 ba59f06.x
C 18	554.2	51.3	554.2	10	BE839427	BE839427 RC3-189754
C 19	550.8	51.0	550.8	12	BI546850	BI546850 603189754
C 20	547	50.6	547	9	AI186701	AI186701 qe82d12.x
C 21	538.2	49.8	538.2	10	BE839460	BE839460 RC3-FN014
C 22	537.8	49.8	537.8	10	BF114684	BF114684 7166c08.x
C 23	528	48.9	528	9	AI800614	AI800614 wgl12e12.x
C 24	521	48.2	521	9	AI458721	AI458721 tk13h10.x
C 25	513	47.5	513	9	AI857399	AI857399 w155f09.x
C 26	513	47.5	513	9	AI885642	AI885642 w159h01.x
C 27	508.4	47.1	508.4	13	BQ206861	BQ206861 UI-R-DZ1-
C 28	507.2	47.0	507.2	13	BU196949	BU196949 AGENCOURT
C 29	496.2	45.9	496.2	12	BI495874	BI495874 df121d04.
C 30	491.4	45.5	491.4	13	BU332089	BU332089 603869689
C 31	482	44.6	482	9	AA746304	AA746304 oa56d03.x
C 32	481.6	44.6	481.6	13	BU225297	BU225297 603799281
C 33	480.2	44.5	480.2	10	BF732679	BF732679 nae05d11.
C 34	477	44.2	477	10	BF513151	BF513151 UI-H-BW1-
C 35	472.2	43.7	472.2	13	BU259874	BU259874 603504254
C 36	469.2	43.4	469.2	10	BE831862	BE831862 RC6-WT006
C 37	468.6	43.4	468.6	9	AA797495	AA797495 vm28b01.x
C 38	466.4	43.2	466.4	10	BG189892	BG189892 RST8946 A
C 39	465.4	43.1	465.4	13	BU613909	BU613909 UI-M-FR0-
C 40	462.8	42.9	462.8	9	AI818729	AI818729 wk91d10.x
C 41	455	42.1	455	9	AM052211	AM052211 wy87d06.x
C 42	454	42.0	454	12	BQ033210	BQ033210 UI-1-CF0-
C 43	452.8	41.9	452.8	9	AI126389	AI126389 qc55e11.x
C 44	447.2	41.4	447.2	9	AA204672	AA204672 zr87c06.x
C 45	446.4	41.3	446.4	14	CA504580	CA504580 UI-R-FS1-

ALIGNMENTS

RESULT 1  
AK048800  
LOCUS  
DEFINITION  
2643 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone: C230071A14 product: inhibitor of growth  
family, member 1-like, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AK048800  
AK048800.1 GI:26339579  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

REFERENCE  
AUTHORS

3  
Kawai, J., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M.,  
Sumi, N., Iguchi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

REFERENCE  
AUTHORS

4  
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Giasi, C., King, B., Koichiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Kochiwa, H.,  
Quackenbush, J., Schraml, L.M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., De Bonaldo, M.F., Brownstein, M.J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schombach, C., Seiya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsuke, S.,  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

REFERENCE  
AUTHORS

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

REFERENCE  
AUTHORS

6 (bases 1 to 2643)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tonaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

REFERENCE  
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
Kanagawa 230-0445, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers

## FEATURES

## source

1..2643  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:C230071A14"  
/db\_xref="taxon:10090"  
/clone="C230071A14"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
476..1321  
/note="unnamed protein product; inhibitor of growth  
family, member 1-like (MGD|MGI:1916510, GB|NM\_023503,  
evidence: BLASTN, 98%, match=971)  
putative"  
/codon\_start=1  
/protein\_id="BAC33461.1"  
/db\_xref="GI:26339580"  
/translation="MLGQQQQQYSSAALLTGERSLLSVQVQYLECVSLPHDQ  
RNVSLRELNDKYQFTLEIDVVEKYKEDDSNOKRLOHLOHRLINSLRDEKI  
QIVTMLVLRNARQMLHSQCQFDPAESRSPKSKWDSQPSRSPRRPRT  
SRDLCHMTNGIDDDDDQPPKRSKSAKKRKSQKAKQREASPFVFAIDPNPTCYLC  
NQSXGEMIGDCNEQCPKEWFFHSCVSLTYPKGWYCPKCRGDNEKTMDSKTEKTK  
ERRAR"  
BASE COUNT 683 a 557 c 710 g 693 t  
ORIGIN  
Query Match 76.7%; Score 828; DB 11; Length 2643;  
Best Local Similarity 88.3%; Pred. No. 1.2e-94;  
Matches 933; Conservative 0; Mismatches 118; Indels 6; Gaps 3;  
QY 1 GCGGCGCGCGCGGTCATGTGCGCTGCTGATCGGAGCGCGCGCGCGCGCGGAT 60  
DB 410 GCGGCGCGCGCGCGGAGGAGCAGCAGCGCGCGCGCGCGCGGCGGCGG 468  
QY 61 CGGCAGGATGTTAGG---GCAGCAGCAGCAGCACTGTACTCGTGGCTCGCTCTGAC 117  
DB 469 CGGCAGGATGTTAGGCGCAGCAGCAGCAGCAGCTGTACTCTGCGCGCGCTCTGAC 528  
QY 118 CGGAGGCGGAGCGCGGTGCTCCTCTGCTACGTGAGGAGTACCTTGAAGTGGTGGAGTC 177  
DB 529 CGGAGGCGGAGCGCGGTGCTCCTCTGCTACGTGAGGAGTACCTTGGAGTGC 598  
QY 178 GCTGCCCGCAGCAGCAGCAGGAGAACTGTCTGCTGCGAGAGCTGGACAAATATCA 237  
DB 589 GCTGCCCGCAGCAGCAGCAGGAGAACTGTCTGCTGCGAGAGCTGGACAAATATCA 648  
QY 238 AGAAACGTTAAAGAAATGATGTCTACGAAAAATATAAGAAAGAGATGATTAAA 297  
DB 649 AGAAACGTTAAAGAAATGATGTCTATGAAAAATATAAGAAAGAGATGATTCAA 708  
QY 298 CCAGAGAAACGCTTACAGAGCTTCTCCAGAGAGCATTAAATATAGTCAAGAAATGGG 357  
DB 709 CCAGAAAAACGCTTACAGAGCAGCTCTCCAGAGAGCGTTAAATCAATAGCAAGATGGG 768  
QY 358 AGATGAAAAATACAGATGTTACACAAATGCTCCAAATGTTGGAAAAATCGGCGAGACA 417  
DB 769 AGATGAAAAATTCAGATGTTTCCAGAGATGTTCCAGATGTTGGAGAAACCGAGCGAGACA 828  
QY 418 AATGAGATTACACTCAGAGTGTTCAGAGATCCTGCTGAAAGTGAACAGCGCTCAGATAA 477  
DB 829 AATGAGCTGCATTACAGTGTTCAGAGATCCTGCTGAAAGTGAACAGCGCTCAGACAA 888  
QY 478 AGCAAGATGGATTCCAGCAACACAGAAAGATCTTCAAGAAAGACCCCGCAGGAGCGGAC 537  
DB 889 GTCAAGATGGATTCCAGTCAACCGGAAAGATCTTCTAGAAGACCTCGAAGACAGAGGAC 948  
QY 538 CAGTGAAGCCGTGATTATGTACATGCGAAATGGATTGAAGTGAAGTGTGATGATCAGCC 597  
DB 949 CAGTGAAGCCGTGATTATGTACATGCGAAATGGATTGAAGTGAAGTGTGATGATCAACC 1008  
QY 598 ACCTAAAGAAAAAATCCAACTCAGCAAAAGAAAGAAACCGCTCCAAAGCCCAAGCAGGA 657

```

Db      1009 ACCGAAAGAAAAGAGAGATCCAAAGTCCGCCAAGAGAAGAAGCGCTCCAAGGCCCAAGCAGGA 1068
QY      658 AAGGGAAGCTTACCTGTTGAGTTTGCAATAGATCCTAATGAACCTACATCTACTGCTTATG 717
Db      1069 GAGGAGGAGCATCCCTGTCGAGTTTGCCATCGATCCCAATGAGCCACCTACTGCTTGTG 1128
QY      718 CAACCAAGTGTCTTATGCGGAGATGATAGATGTGCAATGAACAGTGTCCCAATGAATG 777
Db      1129 TAACCAAGTGTCTTACGCGGAGATGATAGTGTGCAATGAACAGTGTCCCAATGAATG 1188
QY      778 GTTTCACCTTTCATGTTTCACTTACCTATTAACCAAGGGAAGTGTATGCCCCAAA 837
Db      1189 GTTTCACCTTTCATGTTTCACTTACCTATTAACCAAGGGAAGTGTATGCCCCAAA 1248
QY      838 GTGAGGGAGATATATCAGAAAAAACAATGGAACAAAGTACTGAAAGACAAAAAGGATAG 897
Db      1249 GTGTAGGGAGACAATGAGAAACCAATGACAAAGTACCGAAAGACAAAAAGGAGAG 1308
QY      898 AGATCGAGGTAGTAAGGCCATCCACATTTTAAAGGGTATTTGCTTTTATATATTC 957
Db      1309 AAGAGCGAGGTAGTGAAGGCCATC--GGTTTAAAGGGTCTTTGCTTTTATAGAAATG 1366
QY      958 GTTTCCTTTCAGAAATGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAATCA 1017
Db      1367 CTTTACTTTTCAGAAATGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAATCA 1426
QY      1018 TTAGTATTAATGTTGTTATTAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAATCA 1426
Db      1427 TTAGTATTAATGTTGTTATTAAGGTTTAAAGGTTTAAAGGTTTAAATCA 1426

RESULT 2
LOCUS   AK083144
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length
          enriched library, clone: C630016P10 product: inhibitor of growth
          family, member 1-like, full insert sequence.
ACCESSION AK083144
VERSION   AK083144.1 GI:26350286
KEYWORDS HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE    High-efficiency full-length cDNA cloning
JOURNAL  Meth. Enzymol. 303, 19-44 (1999)
MEDLINE  99279253
PUBMED   10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
          Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE    Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL  Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE  20493374
PUBMED   11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
          Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
          Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
          Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
          Fujiwaki, S., Inoue, K., Ito, Y., Iizawa, M., Ohara, E., Watanabe, M.,
          Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
          Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE    RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multicapillary sequencer
JOURNAL  Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE  20530913
PUBMED   11076861
REFERENCE 4
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

```

```

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Mazzionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2812)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. .2812
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C630016P10"
/db_xref="taxon:10090"
/clone="C630016P10"
/sex="male"
/tissue_type="hippocampus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
479. .1324
/note="unnamed protein product; inhibitor of growth
family, member 1-like (MGD|1916510, GB|NM_023503,
evidence: BLASTN, 98%, match=971)"
CDS

```

putative"  
/codon\_start=1  
/protein\_id="BAC38783.1"  
/db\_xref="GI:26350287"  
/translation="MLGQOQOQOQYSSAALLTGERSLRLSCVYQVYLECVESLPDHMQ  
RVSVLRDLNDKYETLKEIDIDVYKEDDSNQKRIQQLQALINSOLIGDEKI  
QIVTQMLELVENRAROMELHSCFODPAESRASDKMSQPSRSPRRPQRTSE  
SRDLCHMTNGIDDCDDPPKEKRSKAKKESKAKOBEREASPVFAIDPNEPTVCLC  
NOVSGEMIGCDNEQCPIEFPHFSCVSLTYPKPKWYCPKCRGDKNDKXSTETKK  
ERRR"  
BASE COUNT 725 a 589 c 743 g 753 t 2 others  
ORIGIN

Query Match 76.6%; Score 827.6; DB 11; Length 2812;  
Best Local Similarity 88.2%; Pred. No. 1.3e-94;  
Matches 932; Conservative 0; Mismatches 119; Indels 6; Gaps 3;  
QY 1 GCGCGCGCGCGCGTGCATGTGCGGCTGCTGATCGGAGCGCGCGCGCGCGGAT 60  
DB 413 GCGCGCGCGCGAGGAGCANCAGCGCGCGCGCGGCGGTGGTG-GCGG 471  
QY 61 CGGCGAGGATTTAGGCGAGCAGCAGCAGCAA---CTGTACTCTCGGCTCGCTCTGAC 117  
DB 472 CGGCGAGGATTTAGGCGAGCAGCAGCAGCAGCTGTACTCTGCGCGCGCTCTGAC 531  
QY 118 CGGCGAGCGGCGCGCTGCTCACCTGCTACGTCAGGAGCTACCTTGAGTCGTCGAGTC 177  
DB 532 CGGAGAGCGGAGCGCGCTGCTCTCTGCTACGTCAGGAGCTACCTGAGTGTGTGAGTC 591  
QY 178 GCTGCCCGCCAGCATCGCAGGAGAACTGTCTGTGCTGCGAGAGCTGCGACAAATATCA 237  
DB 592 GCTGCCCGCCAGCATCGCAGGAGAACTGTCTGTGCTGCGGAGCTGCGACAAATATCA 651  
QY 238 AGAAACGTTAAAGGAAATGATGCTCTACGAAATATAGAAAGAGATGATTTAA 297  
DB 652 AGAAACGTTAAAGGAAATGATGCTCTATGAAATATAGAAAGAGATGATTTCAA 711  
QY 298 CCGAGAGAAACGCTACAGCAGCTTCTCCAGAGCAGCACTAATTAAGTCAAGATGGG 357  
DB 712 CCGAGAGAAACGCTACAGCAGCTTCTCCAGAGAGCGTTAATCAATAGCAGAAATGGG 771  
QY 358 AGATGAAAAAATACAGATGTTTACACAAATGCTCGAAATGGTGAAAAATCGGCAAGACA 417  
DB 772 AGATGAAAAAATACAGATGTTTACACCAAGATGCTCGAATGGTGAGAAACGAGCGAGACA 831  
QY 418 AATGGAGTTACACTACAGTGTTCAGAGATCTCTGAAAGTGAAGAGCGCTTCAGATAA 477  
DB 832 AATGGAGCTGCAATTCACAGTGTTCAGAGATCTCTGAAAGTGAAGAGCGCTTCAGACAA 891  
QY 478 AGCAAGATGGATTCAGCCCAACAGAAAGATCTTCAAGAGACCCCGAGGCGGAGC 537  
DB 892 GTCAAGATGGATTCAGCTCAACCGGAAAGATCTTCTAGAAGACCTCGAAGACAGAGGAC 951  
QY 538 CAGTGAAGCGGTGATTTATGTCACATGGCAATGGATTTGAAGACTGTGATGATCAGCC 597  
DB 952 CAGTGAAGCGGTGATTTATGTCACATGACAAACGGGATGACACTGTGATGATCAACC 1011  
QY 598 ACCTAAAGAAAAAGAAATCCAAAGTCAGCAAGAAAAAGAAACGCTCCCAAGCCCAAGCAGGA 657  
DB 1012 ACCGAAAGAAAAAGAGATCCAAAGTCCGCAAGAAAGAAAGAGCGCTCCCAAGCCCAAGCAGGA 1071  
QY 658 AAGGGAAGCTTACCTGTTGAGTTTGAATAGATCTCTAATGAACCTACATCTGCTTATG 717  
DB 1072 GAGGAGGAGCATCCCTGTCGAGTTTGCATCGATCCCAATGAGCCCACTACTGCTTGTG 1131  
QY 718 CAACCAAGTGTCTTATGGGAGATGATGATGTCGACATGAACAGTGTGCTCAATTCGAATG 777  
DB 1132 TAACCAAGTGTCTTACGAGGAGATGATGATGTCGACATGAACAGTGTGCTTCCATGATG 1191  
QY 778 GTTTCACCTTTTCAATGTTTCACTTACCTATTAACCAAGGGGAAATGGTATTGCCCCAAA 837  
DB 1192 GTTTCACCTTTTCAATGTTTCACTTACCTATTAACCAAGGGGAAATGGTATTGCCCCAAA 1251

QY 838 GTGCAGGGGAGATAATGAGAAAAAATGCAAAAAAGTACTGAAAAAGACAAAAAGGATAG 897  
DB 1252 GTGTAGGGGAGACATGAGAAAAACCATGCAAAAAAGTACCGAAAAAGACAAAAAGGAG 1311  
QY 898 AAGATCGAGGTAGTAAAGCCCATCCACATTTTAAAGGGTATTATTGTCTTTTATATATTC 957  
DB 1312 AAGAGCGAGGTAGTCAAGGCCATC--GGTTTTAAAGGGTCTTCTGCTTTTATAGAAATG 1369  
QY 958 GTTTCCTTTCAGAAAAATGTTTAGGGTAAATGCAATAGCACTATGCAATAATTTTAAATCA 1017  
DB 1370 CTTTACTTTCAGAAAAATGTTTAGGGTAAATGCAATAGCACTATGCAATAATTTTAAATCA 1429  
QY 1018 TTAGTATTAAATCGTCTATTAAAGTGTGTGACTTTG 1054  
DB 1430 TTAGTATTAAATCGTCTATTAAAGTGTGTGACTTTG 1466  
AK012716 1464 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
enriched library, clone:281001M06 product:inhibitor of growth  
family, member 1-like, full insert sequence.  
AK012716  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
PUBMED 10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
PUBMED 11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,  
Sumi, N., Teshi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
PUBMED 11076861  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, W., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,





Db 1366 TGCTTTACTTTTCAGAAAATGTTTTAGGGTAACATGTCATAGAGACTATGCAATAATTTTAAAT 1425  
 QY 1016 CATTAGTATTATGGTGTATTAAAGCTGTGTGACTTTG 1054  
 Db 1426 CATTAGTATTATGGTGTATTAAAGCTGTGTGACTTTG 1464

RESULT 4  
 BQ277444  
 LOCUS  
 DEFINITION BQ277444 .912 bp mRNA linear EST 07-MAY-2002  
 AGENCOURT\_6738254 NIH\_MGC\_127 Homo sapiens cDNA clone IMAGE:5810745  
 5', mRNA sequence.  
 ACCESSION BQ277444  
 VERSION BQ277444.1 GI:20487652  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 912)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs@mail.nih.gov  
 Tissue Procurement: NCI  
 cDNA Library Preparation: Michael Brownstein Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LHC2058 row: h column: 10  
 High quality sequence stop: 587.  
 Location/Qualifiers  
 1. .912  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5810745"  
 /issue\_type="mixed (pool of 40 RNAs)"  
 /lab\_host="DH10B (T1-phage-resistant)"  
 /clone\_lib="NIH\_MGC\_127"  
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);  
 Site 2: SfiI (ggccctccggc); Double-stranded cDNA was  
 prepared from a pool of 40 cell line polyA+ RNAs (bladder  
 - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
 kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
 ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
 gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
 used in cloning as follows:  
 5'-AAGCAGTGGTATCAAGCAGTGGCCATTACGCCGGG-3' and  
 5'-ATTCTAGAGCCGAGCGGCCGATG-3' (30)NN-3'. Full-length  
 enriched library was constructed using the Clontech  
 Creator SMART kit and size-selected to contain the 1-2 kb  
 size fraction (other fractions present in NIH\_MGC\_126 and  
 NIH\_MGC\_128). Library created in the laboratory of T.  
 Udwin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH\_MGC  
 Library."

BASE COUNT 313 a 184 c 235 g 180 t  
 ORIGIN  
 Query Match 69.2%; Score 747.8; DB 13; Length 912;  
 Best Local Similarity 96.9%; Pred. No. 2.2e-84;  
 Matches 794; Conservative 0; Mismatches 22; Indels 3; Gaps 3;  
 QY 47 GCGACGGCGGATCGCAGGATGTAGGACACAGCAGCAGCACTGTCTGGCT 106  
 Db 1 GCGACGGCGGATCGCAGGATGTAGGACACAGCAGCAGCACTGTCTGGCT 60  
 QY 107 GCCTCTCTGACCGGGAGCGGCGCTGCTCACCTGCTACGTCAGGACTACCTTGAG 166

## FEATURES

source

RESULT 5  
 BM982877/c  
 LOCUS  
 DEFINITION UI-CF-EN1-acs-d-05-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone  
 UI-CF-EN1-acs-d-05-0-UI 3', mRNA sequence.  
 ACCESSION BM982877  
 VERSION BM982877.1 GI:19606826  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 735)  
 Ronaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 JOURNAL 97044477  
 MEDLINE 8889548  
 PUBMED  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa

BM982877 735 bp mRNA linear EST 21-FEB-2003  
 UI-CF-EN1-acs-d-05-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone  
 UI-CF-EN1-acs-d-05-0-UI 3', mRNA sequence.  
 ACCESSION BM982877  
 VERSION BM982877.1 GI:19606826  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 735)  
 Ronaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 JOURNAL 97044477  
 MEDLINE 8889548  
 PUBMED  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa

Db 61 GCGCTCTGACCGGGAGCGGAGCGGCTGCTCACCTGCTACGTCAGGACTACCTTGAG 120  
 QY 167 TCGGTGGAGTCTGCTCCACGACATGACAGAGAACTGTCTGTCTGCTGAGAGCTGGAC 226  
 Db 121 TCGGTGGAGTCTGCTCCACGACATGACAGAGAACTGTCTGTCTGCTGAGAGCTGGAC 180  
 QY 227 AACAAATATCAAGAAACGTTAAAGGAAATTTGATGATGTTCTACGAAAAATATAAGAAAGAA 286  
 Db 181 AACAAATATCAAGAAACGTTAAAGGAAATTTGATGATGTTCTACGAAAAATATAAGAAAGAA 240  
 QY 287 GATGATTTAAACACAGAAAGCTCTACAGAGCTTTCCAGAGAGCACTAATTAATAGT 346  
 Db 241 GATGATTTAAACACAGAAAGCTCTACAGAGCTTTCCAGAGAGCACTAATTAATAGT 300  
 QY 347 CAAGAAATTTGGGAGATGAAAAATATACAGATTGTTTACACAAATCTCGAAATTTGGTGAANAAT 406  
 Db 301 CAAGAAATTTGGGAGATGAAAAATATACAGATTGTTTACACAAATCTCGAAATTTGGTGAANAAT 360  
 QY 407 CGGGCAAGACAAATGGAGTTTACACTCACAGTGTTCCTCAAGATCTCTGCTGAAAGTGAACGA 466  
 Db 361 CGGGCAAGACAAATGGAGTTTACACTCACAGTGTTCCTCAAGATCTCTGCTGAAAGTGAACGA 420  
 QY 467 GCCTCAGATAAAGCAAGATGGATTCAGCCAAACAGAAAGATCTTCAAGAAAGACCCCGC 526  
 Db 421 GCCTCAGATAAAGCAAGATGGATTCAGCCAAACAGAAAGATCTTCAAGAAAGACCCCGC 480  
 QY 527 AGGCAAGGACCAAGTGAAGCCGCTGATTTATGTCACATGGCAAAATGGGATTCGAAGACTGT 586  
 Db 481 AGGCAAGGACCAAGTGAAGCCGCTGATTTATGTCACATGGCAAAATGGGATTCGAAGACTGT 540  
 QY 587 GATGATCAGCCACCTTAAAGAAAGAAATCCAAAGTCAGCAAGAAAGAAAGAAAGCTCCAAG 646  
 Db 541 GATGATCAGCCACCTTAAAGAAAGAAATCCAAAGTCAGCAAGAAAGAAAGAAAGCTCCAAG 600  
 QY 647 GCCAAGCAGGAAAGGAGCTTCACCTGTGTGATTTGCAATAGATCTTAATGAACCTACA 706  
 Db 601 GCCAAGCAGGAAAGGAGCTTCACCTGTGTGATTTGCAATAGATCTTAATGAACCTACA 660  
 QY 707 TACTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTCACATGAACAGTGT 766  
 Db 661 TACTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTCACATGAACAGTGT 719  
 QY 767 CCAATTTGAATGTTTTCAC-TTTTTCATGTTTTCACCTTACCTATAAA-CCAAAGGGGAAAT 824  
 Db 720 CCAATTTGAATGTTTTCACCTTTCATGTTTTCACCTTACCTATAAAAGGGGAAAT 779  
 QY 825 GGTATTTGCCCAAGTGCAGGGGAGATATGAGAAAAACA 863  
 Db 780 GGTATTTGCCCAAGTGCAGGGGAGATATGAGAAAAACA 818

RESULT 5  
 BM982877/c  
 LOCUS  
 DEFINITION UI-CF-EN1-acs-d-05-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone  
 UI-CF-EN1-acs-d-05-0-UI 3', mRNA sequence.  
 ACCESSION BM982877  
 VERSION BM982877.1 GI:19606826  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 735)  
 Ronaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 JOURNAL 97044477  
 MEDLINE 8889548  
 PUBMED  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa



```

Query Match      67.4%; Score 727.6; DB 10; Length 798;
Best Local Similarity 98.9%; Pred. No. 7.7e-82;
Matches 785; Conservative 0; Mismatches 4; Indels 5; Gaps 5;

QY 261 ATGTCTACGAAAAATATAGAAAGAGATGATTAAACCCAGAGAAACGCTCTACAGCAGC 320
DB 789 ATGTCTACGCAAAATAT-AGAAGAAGATGATTAAACCCAGAGAAACGCTCTACAGCAGC 731
QY 321 TTCTCCAGAGACACTTAATTAATAGTCAAGAATTTGGAGATGAAAAAATACAGATTGTTA 380
DB 730 TTCTCCAGAGACACT-ATTATAGTCAAGAATTTGGAGATGAAAAA-ACAGATTGTTT- 674
QY 381 CACAATGCTCGAATTTGGTGGAAATTCGGCGAAGCAAAATGAGTTACACTCACAGTGT 440
DB 673 CACAATGCTCGAATTTGGTGG-AAAATCGGCGAAGCAAAATGAGTTACACTCACAGTGT 615
QY 441 TCACAAGATCCTGCTGAAAGTGAACAGAGCTTCAGATAAAGCAAGATGGAATTCAGGCAAC 500
DB 614 TCACAAGATCCTGCTGAAAGTGAACAGAGCTTCAGATAAAGCAAGATGGAATTCAGGCAAC 555
QY 501 CAGAAGATCTTCAAGAAGACCCCGCAGCAGCGGACCAAGTGAAGCCCTGATTTATGTC 560
DB 554 CAGAAGATCTTCAAGAAGACCCCGCAGCAGCGGACCAAGTGAAGCCCTGATTTATGTC 495
QY 561 ACATGGCAATGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAGT 620
DB 494 ACATGGCAATGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAGT 435
QY 621 CAGCAAGAAAAAGAAACCTTCAAGCCCAAGCAGGAAAGGAAAGCTTCACTGTTGAGT 680
DB 434 CAGCAAGAAAAAGAAACCTTCAAGCCCAAGCAGGAAAGGAAAGCTTCACTGTTGAGT 375
QY 681 TTGCAATAGATCCTAATGAACCTTACATCTGTTATGCAACCAAGTGTCTTATGGGAGA 740
DB 374 TTGCAATAGATCCTAATGAACCTTACATCTGTTATGCAACCAAGTGTCTTATGGGAGA 315
QY 741 TGATAGGATGTACATGATGAACAGTGTCCAAATGAATGGTTTTCATTTTCATGTTTTCAC 800
DB 314 TGATAGGATGTACATGATGAACAGTGTCCAAATGAATGGTTTTCATTTTCATGTTTTCAC 255
QY 801 TTACCTTATAAACAAGGGGAATGTTATGCCCAGAGTGCAGGGGAGATATAGAAAA 860
DB 254 TTACCTTATAAACAAGGGGAATGTTATGCCCAGAGTGCAGGGGAGATATAGAAAA 195
QY 861 CAATGGACAAAAGTACTGAAAAAGACAAAAAAGSATAAGAGATCGAGGTAGTAAAGCCAT 920
DB 194 CAATGGACAAAAGTACTGAAAAAGACAAAAAAGSATAAGAGATCGAGGTAGTAAAGCCAT 135
QY 921 CCACATTTTAAAGGGTTATTTGCTTTTATATAATTCGTTTTCGTTTCAGAAAAATGTTTA 980
DB 134 CCACATTTTAAAGGGTTATTTGCTTTTATATAATTCGTTTTCGTTTTCAGAAAAATGTTTA 75
QY 981 GGGTAAATCATAAGACTATGCAATATTTTAACTATTAGTATTAAATGCTGTATTAAAA 1040
DB 74 GGGTAAATCATAAGACTATGCAATATTTTAACTATTAGTATTAAATGCTGTATTAAAA 15
QY 1041 GTTGTGTACTTTG 1054
DB 14 GTTGTGTACTTTG 1

```

```

RESULT 7
BE796780
LOCUS 601587557E1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3941655 5',
DEFINITION mRNA sequence.
ACCESSION BE796780
VERSION BE796780.1 GI:10218080
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

# REFERENCE

## AUTHORS

### TITLE

#### JOURNAL

##### COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 793)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue procurement: DCTD/DTF  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: L10W794 row: a column: 16  
 High quality sequence stop: 769.  
 Location/Qualifiers  
 1..793

# FEATURES

## source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3941655"  
 /tissue\_type="small cell carcinoma"  
 /cell\_line="MGC3"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 7"  
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCAAGGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 240 a 177 c 230 g 146 t

# Query Match

## Best Local Similarity

### Matches

55.8%; Score 711; DB 10; Length 793;  
 98.9%; Pred. No. 9.1e-80;  
 747; Conservative 0; Mismatches 5; Indels 3; Gaps 3;  
 QY 1 GCGGCGCGCGCGGTCATGTGCGGCTGCTGATGCGAGGCGGCGGCGCGCGGAT 60  
 DB 41 GCGGCGCGCGCGGTCATGTGCGGCTGCTGATGCGAGGCGGCGGCGCGGAT 100  
 QY 61 CGCGAGGATGTTAGGCGAGCAGCAGCAACTGTACTGCTGCGCTGCGCTCTCTGACCG 120  
 DB 101 CGCGAGGATGTTAGGCGAGCAGCAGCAACTGTACTGCTGCGCTGCGCTCTCTGACCG 160  
 QY 121 GGAGCGGAGCGGCTGCTCACCTGCTACGTGAGGACTACCTTGTAGTGGTGGAGTCGCT 180  
 DB 161 GGAGCGGAGCGGCTGCTCACCTGCTACGTGAGGACTACCTTGTAGTGGTGGAGTCGCT 220  
 QY 181 GCCCCACGACATGCGAGAGAACGTCGTCTGTGCTGCGAGAGCTGGACCAACAATATCAAGA 240  
 DB 221 GCCCCACGACATGCGAGAGAACGTCGTCTGTGCTGCGAGAGCTGGACCAACAATATCAAGA 280  
 QY 241 AACGTTAAAGGAAATTTGATGTCCTACGAAAAATATAAGAAAGAGATGATTAAACCA 300  
 DB 281 AACGTTAAAGGAAATTTGATGTCCTACGAAAAATATAAGAAAGAGATGATTAAACCA 340  
 QY 301 GAAGAAACGCTACAGCAGCTTCTCCAGAGCAGCTAAATTAATAGTCAAGAATTGGGAGA 360  
 DB 341 GAAGAAACGCTACAGCAGCTTCTCCAGAGCAGCTAAATTAATAGTCAAGAATTGGGAGA 400  
 QY 361 TGAATAAATACAGATTGTTTACACAAATGCTCGAATTTGGTGGAAAAATCGGCGAGACAAAT 420  
 DB 401 TGAATAAATACAGATTGTTTACACAAATGCTCGAATTTGGTGGAAAAATCGGCGAGACAAAT 460  
 QY 421 GGAGTTACACTCACAGTGTTCACAGATCCTCTGAAAGTGAACGAGACCTCAGATAAAGC 480  
 DB 461 GGAGTTACACTCACAGTGTTCACAGATCCTCTGAAAGTGAACGAGACCTCAGATAAAGC 520  
 QY 481 AAAGATGGATTCCAGCCACAGAAAGATCTTCAAGAAAGACCCCGCAGGCGGACCGAG 540

Db 521 AAGATGGATTCCAGCCCAACAGAAAGATCTTCAAGAAGACCCCGCAGGCGAGCCAG 580  
 QY 541 TGAAGCCGTGATTTA-TGTACATGCGCAATGGGATTGAAGACTGTGTGATGATCAGCCAC 599  
 Db 581 TGAAGCCGTGATTTATTGTTCATGCGCAATGGGATTGAAGACTGTGTGATGATCAGCCAC 640  
 QY 600 CTTAAGAAAGAAATCCAACTGAGCAAGAAAGAAACGCTCAAGGCCCAAGCAGAAA 659  
 Db 641 CTTAAGAAAGAAATCCAACTGAGCAAGAAAGAAACGCTCAAGGCCCAAGCAGAAA 700  
 QY 660 GGAAGCTTCACCTGTGAGTTTGCATAGATCCTTAATGAACCTACATATCTTATGCA 719  
 Db 701 GGAAGCTTCACCTGTGAG-TTGCATAGATTTCTAATGAACCTACATATCTTATGCA 759  
 QY 720 ACCAAGTGTCTTATGCGGAGATGATAGGATGTGAC 754  
 Db 760 ACCAAGTGTCTT-GGGGAGATGATACGGTGTGCC 793

RESULT 8  
 BI548536  
 LOCUS  
 DEFINITION  
 60319125F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5262540 5',  
 mRNA sequence.  
 ACCESSION  
 VERSION  
 BI548536  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 778)  
 NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL  
 Unpublished  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-i@mail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Pietro Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LLM11661 row: n column: 13  
 High quality sequence stop: 775.  
 Location/Qualifiers  
 1..778  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5262540"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH108"  
 /clone\_lib="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 237 a 176 c 227 g 138 t

## FEATURES

source  
 1..778  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5262540"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH108"  
 /clone\_lib="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 237 a 176 c 227 g 138 t

## Query Match

Best Local Similarity 64.3%; Score 694.2; DB 12; Length 778;  
 Matches 718; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 GCGGCGGCGCGGCTGCATGTGCGGCTGCTGATGCGAGGCGGCGCGGCGGAT 60  
 Db 57 GCGGCGGCGGCGGCTGCATGTGCGGCTGCTGATGCGAGGCGGCGGCGGAT 116  
 QY 61 CGGCGAGGATTTAGGCGCAGCAGCAGCAACTGTACTCGTGGCTCGGCTCGTGGCGG 120  
 Db 117 CGGCGAGGATTTAGGCGCAGCAGCAGCAACTGTACTCGTGGCTCGGCTCGTGGCGG 176  
 QY 121 GGAAGCGGAGCCGCTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 Db 177 GGAAGCGGAGCCGCTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 236  
 QY 181 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 240  
 Db 237 GCCACACGACATGCGAGGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 296  
 QY 241 AACGTTAAAGGAAATGATGATGCTTACGAAAATATAAGAAAGATGATTTAAACCA 300  
 Db 297 AACGTTAAAGGAAATGATGATGCTTACGAAAATATAAGAAAGATGATTTAAACCA 356  
 QY 301 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCAGCTAAATTAATAGTCAAGAAATGGGAGA 360  
 Db 357 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCAGCTAAATTAATAGTCAAGAAATGGGAGA 416  
 QY 361 TGAAGAAATACAGATTGTTTACACAAATGCTCGAAATTTGGTGGAAATCGGSCAAGACAAAT 420  
 Db 417 TGAAGAAATACAGATTGTTTACACAAATGCTCGAAATTTGGTGGAAATCGGSCAAGACAAAT 476  
 QY 421 GGAATTACACTCAGCTGTTTCCAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 Db 477 GGAATTACACTCAGCTGTTTCCAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 536  
 QY 481 AAAGATGGAATCCAGCCACCAAGAGATCTTCAAGAACCCCGCAGGCGGCGGACCCAG 540  
 Db 537 AAAGATGGAATCCAGCCACCAAGAGATCTTCAAGAACCCCGCAGGCGGCGGACCCAG 596  
 QY 541 TGAAGCCGTGATTTATGTGCATGCGCAATGGGATTTGAAGATGTGTGATGATGATGATGATGATGAT 600  
 Db 597 TGAAGCCGTGATTTATGTGCATGCGCAATGGGATTTGAAGATGTGTGATGATGATGATGATGATGAT 656  
 QY 601 TAAAGAAAGAAATCCAACTGAGCAAGAAAGAAACGCTCCAAAGCCCAAGCAGGAAAG 660  
 Db 657 TAAAGAAAGAAATCCAACTGAGCAAGAAAGAAACGCTCCAAAGCCCAAGCAGGAAAG 715  
 QY 661 GGAAGCTTCACCTGTGAGTTTGCATAGATCCTATGAACTTACATCTGCTTATGCAA 720  
 Db 716 GGAAGCTTCACCTGTGAGTTTGC-ATAGATCCTAATGAACCTTACATCTGCTTATGCAA 774  
 QY 721 CCA 723  
 Db 775 CAA 777

RESULT 9  
 BG211544  
 LOCUS  
 DEFINITION  
 RST31111 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 ACCESSION  
 BG211544  
 VERSION  
 BG211544.1 GI:13733231  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 919)  
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
 Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,  
 Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith  
 ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher  
 ,J., Danzig,J. and Ducar,M.  
 Creation of genome-wide protein expression libraries using random  
 activation of gene expression  
 Nat. Biotechnol. 19 (5), 440-445 (2001)

```

MEDLINE      21227151
PUBMED       11329013
COMMENT      Contact: Scott J. Cain
              Athersys, Inc.
              3201 Carnegie Ave, Cleveland, OH 44115, USA
              Tel: 216 431 9900
              Fax: 216 361 9596
              Email: scain@atherys.com
              High quality sequence stop: 439.
              Location/Qualifiers
                1..919
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone_lib="HT1080"
                /notes="See 'Creation of Genome-wide Protein Expression
                Libraries using Random Activation of Gene Expression',
                Nature Biotechnology, in press. Note that even though the
                cell type indicated is HT1080, since a random activation
                method was used, these sequence tags are not necessarily
                expressed in HT1080 under normal circumstances."
              BASE COUNT      332 a 151 c 216 g 216 t 4 others
              ORIGIN
                Query Match      63.0%; Score 680.8; DB 10: Length 919;
                Best Local Similarity 95.1%; Pred. No. 4.9e-76;
                Matches 788; Conservative 0; Mismatches 30; Indels 11; Gaps 8;

QY      236  CAAAGAACGTTAAAGGAAATGTATGATGCTCAGCAAAAATATAAGAAAGAGATGATTAA 295
DB      1    CCAAGAACGTTAAAGGAAATGTATGATGCTCAGCAAAAATATAAGAAAGAGATGATTAA 60

QY      296  AACCAGAAGAACGCTCTACAGAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAAATTG 355
DB      61  AACCAGAAGAACGCTCTACAGAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAAATTG 120

QY      356  GAGATGAAATAATACAGATTGTTACACAATCTCGAATTCGTGGAAATCGGCAAGA 415
DB      121  GGAGATGAAATAATACAGATTGTTACACAATCTCGAATTCGTGGAAATCGGCAAGA 180

QY      416  CAAATGGAGTTACACTCACAGTGTCTTCCAGAGTCTCTGCTCAAGAGTCAACGAGCTCAGAT 475
DB      181  CAAATGGAGTTACACTCACAGTGTCTTCCAGAGTCTCTGCTCAAGAGTCAACGAGCTCAGAT 240

QY      476  AAAGCAAAGATGGAATCCAGCCCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGGCGG 535
DB      241  AAAGCAAAGATGGAATCCAGCCCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGGCGG 300

QY      536  ACCAGTGAAGCCGTGATTTATGTCATCGGCAATGGGATTCGAAGACTGTGATGATCAG 595
DB      301  ACCAGTGAAGCCGTGATTTATGTCATCGGCAATGGGATTCGAAGACTGTGATGATCAG 360

QY      596  CCACCTTAAGAAAGAAATCCAAAGTCAAGCAAGAAAGAAAGCGCTCAAGGCGCAAGCAG 655
DB      361  CCACCTTAAGAAAGAAATCCAAAGTCAAGCAAGAAAGAAAGCGCTCAAGGCGCAAGCAG 420

QY      656  GAAAGGGAAGCTTCACCTGTGTGATTTGCAATAGATCTTAATGAACCTCATACTGCTTA 715
DB      421  GAAAGGGAAGCTTCACCTGTGTGATTTGCAATAGATCTTAATGAACCTCATACTGCTTA 480

QY      716  TGAACCAAGTGTCTTATGGGGAGATGATAGGATGTGACAATGAACAGTGTCCAATTGAA 775
DB      481  TGAACCAAGTGTCTTATGGGGAGATGATAGGATGTGACAATGAACAGTGTCCAATTGAA 540

QY      776  TGGTTTCACCTTTTCATGTGTTTCACCTTACCTATAAACCAAGGGAATGGTATGCCCA 835
DB      541  TGGTTTCACCTTTTCATGTGTTTCACCTTACCTATAAACCAAGGGAATGGTATGCCCA 600

QY      836  AAGTCAGGCGGAGATAATGAGAAACAAATGGACAAAGTACTGAAAGACAAAGAGGAT 895
DB      601  NAGTCAGGCGGAGATAATGAGAAACAAAG--ACAAAGCACTGAAAGACAAAGAGG-T 657

```

---

```

QY      896  AGAAGATCGAGGTAGTAAGGCCATCCACATTTAAAGGGTTATTTGTCTTTTATATAAT 955
DB      658  AGAAGATCGAAGTAGTAA--GGCCATCACATTTAAAGGGTATTGTGCTTTATATAAT 714

QY      956  TCGTTTTCCTTTCAGAAAATGTTTATAGGTAATGCAATAGACTATGCAATATTTTAAAT 1015
DB      715  TCGATTTCCTTTCAGAAAATG--TTAGGGTAAATGCTTATAGACTATGCA--TAATNTTAAAT 770

QY      1016  CATTAGTATTAAATGGTGTATTAAAGAGTTGTTGCTTCTTTCAGAAAAAATAA 1064
DB      771  CATTAGTATTAAATGG--GGATAAAGTTGTTGCTTTCAGAAAAAATAA 818

RESULT 10
LOCUS   BM682789/c
DEFINITION UI-E-EJ1-ajf-g-22-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
          UI-E-EJ1-ajf-g-22-0-UI 3', mRNA sequence.
ACCESSION BM682789
VERSION   BM682789.1 GI:18992685
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 675)
AUTHORS   Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
          Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Hominidae; Homo.
TITLE     Bonaldo,M.F., Lennon,G. and Soares,M.B.
          Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@uiowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Researchers may obtain clones from Research
          Genetics (www.resgen.com).
          Seq primer: M13 Forward
          POLYA=Yes.

FEATURES             Location/Qualifiers
     1..675
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone_lib="UI-E-EJ1-ajf-g-22-0-UI"
     /tissue_type="fetal eyes, lens, eye anterior segment,
     optic nerve, retina, Retina Foveal and Macular, RPE and
     Choroid"
     /dev_stage="fetal and adult"
     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
     /clone_lib="UI-E-EJ1"
     /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
     modified polylinker; Site 1: EcoR I; Site 2: Not I;
     UI-E-EJ1 is a subcloned cDNA library constructed
     according to Bonaldo, Lennon and Soares, Genome Research,
     6:791-806, 1996. First strand cDNA synthesis was primed
     with an oligo-dT primer containing a Not I site. Double
     stranded cDNA was ligated to an EcoR I adaptor, digested
     with Not I, and cloned directionally into pT73-Pac
     vector. The oligonucleotide used to prime the synthesis of
     first-strand cDNA contains a library tag sequence that is
     located between the Not I site and the (dT)18 tail. The
     sequence tags for this library are: fetal eyes, AGAATCAAGA
     ; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;
     optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and

```

Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG LIB=UI-E-EJ1  
TAG TISSUE=Retina and Macular Retina  
TAG\_SEQ=GTCC

BASE COUNT 164 a 148 c 110 g 253 t  
ORIGIN

Query Match 62.2%; Score 671.8; DB 12; Length 675;  
Best Local Similarity 99.7%; Pred. No. 7.8e-75;  
Matches 673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 398 GTGGAATCGGCAAGACAAATGGAGTTACACTCAGAGTGTTCCTCAAGATCCTGCTGAA 457  
DB 675 GTGGAATCGGCAAGACAAATGGAGTTACACTCAGAGTGTTCCTCAAGATCCTGCTGAA 616  
QY 458 AGTGAACGACCTCAGATAAGCAAGATGGATTCCGCCAACAGAAAGATCTTCAAGA 517  
DB 615 AGTGAACGACCTCAGATAAGCAAGATGGATTCCGCCAACAGAAAGATCTTCAAGA 556  
QY 518 AGACCCCGCAGGCGGACGAGTGAAGCCGTTGATTATGTCACATGGCAATGGGATT 577  
DB 555 AGACCCCGCAGGCGGACGAGTGAAGCCGTTGATTATGTCACATGGCAATGGGATT 596  
QY 578 GAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAGTCAGCAAGAAAGAAA 637  
DB 495 GAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAGTCAGCAAGAAAGAAA 436  
QY 638 CGCTCAAGCCCAAGCAAGAAAGGAGCTTCACTGTTGAGTTGCAATAGATCTTAAT 697  
DB 435 CGCTCAAGCCCAAGCAAGAAAGGAGCTTCACTGTTGAGTTGCAATAGATCTTAAT 376  
QY 698 GRACCTACATCTGCTTATGCAACCAAGTCTTATGCGGAGATGATAGATGTGCAAT 757  
DB 375 GRACCTACATCTGCTTATGCAACCAAGTCTTATGCGGAGATGATAGATGTGCAAT 316  
QY 758 GAACAGTGTCCAATGGAATGTTTCACTTTCATGTTGTTTCACTTACCTATAACCAAG 817  
DB 315 GAACAGTGTCCAATGGAATGTTTCACTTTCATGTTGTTTCACTTACCTATAACCAAG 256  
QY 818 GGAATATGTTATGCCCCAAGTCAGGGAGATAATGAGAAACAAATGGACAAAGTACT 877  
DB 255 GGAATATGTTATGCCCCAAGTCAGGGAGATAATGAGAAACAAATGGACAAAGTACT 196  
QY 878 GAAAGACAAAGAGATAGAGATCGAGTAGTAAGGCTCCATTTTAAAGGTT 937  
DB 195 GAAAGACAAAGAGATAGAGATCGAGTAGTAAGGCTCCATTTTAAAGGTT 136  
QY 938 ATTGTCTTTTATATATTCGTTTTCAGAAATGTTTATAGGTTAAATGATTAAGAC 997  
DB 135 ATTGTCTTTTATATATTCGTTTTCAGAAATGTTTATAGGTTAAATGATTAAGAC 76  
QY 998 TATGCAATATTTTATATATGTTTATGTTATGTTATGTTATGTTTACTTTGAAA 1057  
DB 75 TATGCAATATTTTATATATGTTTATGTTATGTTATGTTATGTTTACTTTGAAA 16  
QY 1058 AAAAAAAAAAAAAA 1072  
DB 15 AAAAAAAAAAAAAA 1

RESULT 11  
BM719460  
LOCUS  
DEFINITION  
UI-E-EJ1-af-f-g-22-0-UI.r1 664 bp mRNA linear EST 01-MAR-2002  
UI-E-EJ1-af-f-g-22-0-UI 5', mRNA sequence.  
BM719460  
VERSION  
BM719460.1 GI:19038113  
KEYWORDS  
Homo sapiens (human)  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 664)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Prepared by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.regen.com).  
Seq primer: M13 Reverse.

FEATURES  
source

Location/Qualifiers  
1..664  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ1-af-f-g-22-0-UI"  
/tissue\_type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina Foveal and Macular, RPE and  
Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-EJ1"  
/note="Organ: eye; Vector: p7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-EJ1 is a subtracted cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into p7T3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are: fetal eyes, AGAATCAAGA  
; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;  
optic nerve, CCATTAGTGC; retina, CCGCG; Retina Foveal and  
Macular, GTCC; RPE and Choroid, ACCTA. This library was  
created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."  
BASE COUNT 241 a 111 c 149 g 162 t 1 others  
ORIGIN

Query Match 61.4%; Score 663; DB 12; Length 664;  
Best Local Similarity 99.8%; Pred. No. 9.8e-74;  
Matches 663; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 377 GTTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGACAAATGGAGTTTACTACAG 436  
DB 1 GTTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGACAAATGGAGTTTACTACAG 60  
QY 437 TGTTCACAGATCTCTCTGAAAGTGAACGAGCTCAGATAAGCAAGATGGATTCAGC 496  
DB 61 TGTTCACAGATCTCTCTGAAAGTGAACGAGCTCAGATAAGCAAGATGGATTCAGC 120  
QY 497 CAACAGAAAGATCTTCAAGAGAGACCCCGCAGCGACGACGAGTGAAGCCGCTGATTTA 556  
DB 121 CAACAGAAAGATCTTCAAGAGAGACCCCGCAGCGACGACGAGTGAAGCCGCTGATTTA 180  
QY 557 TGTTCACATGGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCC 616  
DB 181 TGTTCACATGGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCC 240

QY 617 AAGTCAGCAAGAAAGAAACGCTCAAGGCCAAGCAGGAAGGAAAGGAGCTTCACTGTT 676  
 |||||  
 Db 241 AAGTCAGCAAGAAAGAAACGCTCAAGGCCAAGCAGGAAGGAAAGGAGCTTCACTGTT 300  
 |||||  
 QY 677 GAGTTTGCAATAGATCCTAATGAACCTACATCTGCTTATGCAACCAAGTGCTTATGGG 736  
 |||||  
 Db 301 GAGTTTGCAATAGATCCTAATGAACCTACATCTGCTTATGCAACCAAGTGCTTATGGG 360  
 |||||  
 QY 737 GAGATGATGAGTGTGACAAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCATGTT 796  
 |||||  
 Db 361 GAGATGATGAGTGTGACAAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCATGTT 420  
 |||||  
 QY 797 TCACCTTACCTATAAACCAAGGGGAAATGTTATGCTCCCAAGTGCGAGGAGAGATAATGAG 856  
 |||||  
 Db 421 TCACCTTACCTATAAACCAAGGGGAAATGTTATGCTCCCAAGTGCGAGGAGAGATAATGAG 480  
 |||||  
 QY 857 AAAACAATGACAAAAGTACTGAAAAGACAAAAAGGATAGAAAGTCGAGGTAGTAAAGG 916  
 |||||  
 Db 481 AAAACAATGACAAAAGTACTGAAAAGACAAAAAGGATAGAAAGTCGAGGTAGTAAAGG 540  
 |||||  
 QY 917 CCATCCACATTTTAAAGGGTTATTTGCTTTTATATAATTCGTTTTCGTTTTCAGAAAATGT 976  
 |||||  
 Db 541 CCATCCACATTTTAAAGGGTTATTTGCTTTTATATAATTCGTTTTCGTTTTCAGAAAATGT 600  
 |||||  
 QY 977 TTTAGGGTAAATGCATAAGACTATGCAATAATTTTAAATCAATAGTATTAAATGCTGTT 1036  
 |||||  
 Db 601 TTTAGGGTAAATGCATAAGACTATGCAATAATTTTAAATCAATAGTATTAAATGCTGTT 660  
 |||||  
 QY 1037 AAAA 1040  
 |||||  
 Db 661 AAAA 664  
 |||||

RESULT 12  
 BI495873/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 HOMO sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 653)  
 Roberton,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bisber,F.R.  
 and Morton,C.C.  
 TITLE  
 Isolation of novel and known genes from a human fetal cochlear cDNA  
 library using subtractive hybridization and differential screening  
 Genomics 23, 42-50 (1994)  
 95130111  
 7829101  
 Contact: Morton, C. C.  
 Departments of Pathology and Obstetrics, Gynecology and  
 Reproductive Biology  
 Brigham and Women's Hospital  
 75 Francis Street, Harvard Medical School, Boston, MA 02115, USA  
 Tel: 617 732 7980  
 Fax: 617 738 6996  
 Email: cmorton@rics.bwh.harvard.edu  
 DNA sequencing and analyses were performed by National Institutes  
 of Health Intramural Sequencing Center (NISC; see  
 http://www.nisc.nih.gov).  
 This clone is available royalty-free through LILNI; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Plats: LLAM6327 row: G column: 7  
 Seq primer: 17 primer.  
 Location/Qualifiers  
 1. .653  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"

## FEATURES

## source

RESULT 13  
 BI091047/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS

BI091047  
 602855235F1 NIH\_MGC\_10 Homo sapiens cDNA clone IMAGE:496593 5',  
 mRNA sequence.  
 BI091047  
 BI091047.1 GI:14509377  
 EST.

/db\_xref="taxon:9606"  
 /clone="IMAGE:2540190"  
 /tissue\_type="cochlea"  
 /dev\_stage="16-22 week fetus"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Morton Fetal Cochlea"  
 /note="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI;  
 Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned  
 unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.  
 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP  
 XR Vector. Library constructed by N. Robertson, C. Morton.  
 -5' adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor  
 sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 160 a 146 c 106 g 241 t

Query Match 58.9%; Score 636.4; DB 12; Length 653;  
 Best Local Similarity 99.8%; Pred. No. 2.1e-70;  
 Matches 637; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 435 AGTGTTCACAGATCCTGCTGAAGTGAACGAGCCCTCAGATAAGCAGAGATGGATTCCA 494  
 |||||  
 Db 653 AGTGTTCACAGATCCTGCTGAAGTGAACGAGCCCTCAGATAAGCAGAGATGGATTCCA 594  
 |||||  
 QY 495 GCCAACCCAGAAAGATCTTCAAGAGACCCCGCAGGCGAGCCAGTGAAGCCGCGATT 554  
 |||||  
 Db 593 GCCAACCCAGAAAGATCTTCAAGAGACCCCGCAGGCGAGCCAGTGAAGCCGCGATT 534  
 |||||  
 QY 555 TATGTCAATGCGAAATGGGATTGAAGACTGTGATGATCAGCCACTTAAAGAAAAGAAAT 614  
 |||||  
 Db 533 TATGTCAATGCGAAATGGGATTGAAGACTGTGATGATCAGCCACTTAAAGAAAAGAAAT 474  
 |||||  
 QY 615 CCAAGTCAGCAAGAAAGAAAGAACGCTCCAGCCCAAGCAGGAGGAGGAGCTTCCACCTG 674  
 |||||  
 Db 473 CCAAGTCAGCAAGAAAGAAAGAACGCTCCAGCCCAAGCAGGAGGAGGAGCTTCCACCTG 414  
 |||||  
 QY 675 TTGAGTTTGAATAGATCCTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTATG 734  
 |||||  
 Db 413 TTGAGTTTGAATAGATCCTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTATG 354  
 |||||  
 QY 735 GGGAGATGATGAGTGTGCAATGAACAGTGTCCAATGGAATGGTTTTCATCTTTTCATG 794  
 |||||  
 Db 353 GGGAGATGATGAGTGTGCAATGAACAGTGTCCAATGGAATGGTTTTCATCTTTTCATG 294  
 |||||  
 QY 795 TTTCTACTTACCTATAAACCAGAGGGAATGTTGCCCAAGTCCAGGAGAGATAATG 854  
 |||||  
 Db 293 TTTCTACTTACCTATAAACCAGAGGGAATGTTGCCCAAGTCCAGGAGAGATAATG 234  
 |||||  
 QY 855 AGAAACCAATGCAACAAAGTACTGAAAAGACAAAAAGGATAGAGATCGAGGTAGTAAA 914  
 |||||  
 Db 233 AGAAACCAATGCAACAAAGTACTGAAAAGACAAAAAGGATAGAGATCGAGGTAGTAAA 174  
 |||||  
 QY 915 GGCCATCCCATTTTAAAGGGTTATTTGCTTTTATATAATTCGTTTTCAGAAAAT 974  
 |||||  
 Db 173 GGCCATCCCATTTTAAAGGGTTATTTGCTTTTATATAATTCGTTTTCAGAAAAT 114  
 |||||  
 QY 975 GTTTAGGGTAAATGCATAAGACTATGCAATAATTTTAAATCAATAGTATTAAATGTTA 1034  
 |||||  
 Db 113 GTTTAGGGTAAATGCATAAGACTATGCAATAATTTTAAATCAATAGTATTAAATGTTA 54  
 |||||  
 QY 1035 TTAAGAGTTGTTGTTACTTTTGAAGAAAAAAGAAAAA 1072  
 |||||  
 Db 53 TTAAGAGTTGTTGTTACTTTTGAAGAAAAAAGAAAAA 16  
 |||||



```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 891)
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE       NIH-MGC http://mgi.nci.nih.gov/
JOURNAL     National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1022 row: e column: 10
            High quality sequence stop: 722.
FEATURES    Location/Qualifiers
             1..891
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4996593"
                /cell_line="MGC36"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_10"
                /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 1.5 kb. Library prepared by Life
                Technologies."
BASE COUNT  237 a 199 c 149 g 306 t
ORIGIN

Query Match      58.5%; Score 632.2; DB 12; Length 891;
Best Local Similarity 96.5%; Pred. No. 5.8e-70;
Matches 687; Conservative 0; Mismatches 21; Indels 4; Gaps 4;

QY 336 TAATTAATAGTCAAGATTCGGAGATGAAATACAGA-TTGTACACAATGCTCGAA 394
Db 711 TAATTAATAGTCAAGATTCGGAGATGAAATACAGA-TTGTACACAATGCTCGAA 652
QY 395 TTGGTGAATACGGCAGACAAATGGAGTTACACTCACAGTGTTCACAGATCCTG-C 453
Db 651 TTGGTGTAAATC-GGCAAGACATATGGAGTTACACTCACAGTGTTCACAGATCCTGTC 593
QY 454 TGAAGTGAACGAGCTTCAGATAAGCAAGAT-GGATTCAGCCACCAAGAAAGATCTT 512
Db 592 TGAAGTGAACGAGCTTCAGATAAGCAAGATGGGATTCAGCCACCAAGAAAGATCTT 533
QY 513 CAAGAGACCCCGCAGCGGACGAGTGAAGCCGTGATTTATGTCATCATGGCAATG 572
Db 532 CAAGAGACCCCGCAGCGGACGAGTGAAGCCGTGATTTATGTCATCATGGCAATG 473
QY 573 GGATTGAAGACTGTGATGATCAGCCACTAAAGAAAGAAATCAAGTCAGCAAGAAAGAA 632
Db 472 GGATTGAAGACTGTGATGATCAGCCACTAAAGAAAGAAATCAAGTCAGCAAGAAAGAA 413
QY 633 AGAAACGCTCAAGGCCAAGCAGGAAAGGGAAGCTTCACCTGTTGAGTTTGCATATGATC 692
Db 412 AGAAACGCTCAAGGCCAAGCAGGAAAGGGAAGCTTCACCTGTTGAGTTTGCATATGATC 353
QY 693 CTAATGACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTG 752
Db 352 CTAATGACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTG 293
QY 753 ACAATGAACAGTGTCCAAATGATGTTTCACCTTTTCATGTTTTCACCTTACCTATAAC 812
Db 292 ACACTGAACAGTGTCCAAATGATGTTTTCACCTTTTCATGTTTTCACCTTACCTATAAC 233
QY 813 CAAAGGGGAATGGTATTTGCCAAAGTCAGGGGAGATAATGAGAAACAAATGGAACAAA 872

```

---

```

Db 232 CAAAGGGGAATGGTATTTGCCAAAGTCAGGGGAGATATGAGAAACAAATGGACAAA 173
QY 873 GTACTGAAAGACAAAAGAGATAGAGATCGAGGTAGTAGTAAGCCATCACAATTTTAAA 932
Db 172 GTACTGAAAGACAAAAGAGATAGAGATCGAGGTAGTAGTAAGCCATCACAATTTTAAA 113
QY 933 GGGTTATTTGCTCTTTTATATATATTCGTTGCTTTCAGAAAATGTTTTCAGGTAATGCAT 992
Db 112 GGGTTATTTGCTCTTTTATATATATTCGTTGCTTTCAGAAAATGTTTTCAGGTAATGCAT 53
QY 993 AGACTATGCCAATAATTTTATCATAGTATTAATGCTGCTATTAAGTTG 1044
Db 52 AGACTATGCCAATAATTTTATCATAGTATTAATGCTGCTATTAAGTTG 1

RESULT 14
LOCUS      BI460319
DEFINITION 603201967F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5267729 5',
            mRNA sequence.
ACCESSION  BI460319
VERSION     BI460319.1 GI:15250975
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 708)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11675 row: f column: 18
            High quality sequence stop: 700.
FEATURES    Location/Qualifiers
             1..708
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5267729"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_97"
                /note="Organ: testis; Vector: pBluescriptR (modified
                pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
                ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
                size-selected for average insert size 2.2 kb and
                normalized to 10^5. This is a primary library enriched
                for full-length clones and constructed using the
                Cap-trapper method (Carninci, in preparation). Library
                constructed by M. Brownstein (NIH/NHGRI, National
                Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT  209 a 163 c 213 g 122 t 1 others
ORIGIN

Query Match      56.2%; Score 607; DB 12; Length 708;
Best Local Similarity 98.9%; Pred. No. 9.2e-67;
Matches 621; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 10 GCCTGTCATGTCGGCTCTGTCGATCGGAGCGGCGGACCGCGGATCGCAGGAT 69
Db 81 GCGCGTGCATGTCGGCTCTGTCGATCGGAGCGGCGGACCGCGGATCGCAGGAT 140
QY 70 GTTAGGGCAGCAGCAGCACTGTACTGTCGCTCGCTCTCTGACCCGGGAGCGGAG 129

```



141 GTTAGGGGAGCAGCAGCAGCACTGTACTCGTCGGCGCGCTCCTGACCGGGAGCGGAG 200  
130 CCGGCTGCTCAGCTGCTAGTGCAGAGCTACCTTGTAGTGGTGGAGTCGTCGCCACGA 189  
201 CCGGCTGCTCAGCTGCTAGTGCAGAGCTACCTTGTAGTGGTGGAGTCGTCGCCACGA 260  
190 CATGAGAGGAGCAGTGTCTGCTGCTGCGAGAGCTGGACCAACAATATCAAGAAACGTTAAA 249  
261 CATGAGAGGAGCAGTGTCTGCTGCTGCGAGAGCTGGACCAACAATATCAAGAAACGTTAAA 320  
250 GGAATTTGATGATGTCTTACGAAATAATATAGAAAGAGATGATTTAAACCCAGAGAAACG 309  
321 GGAATTTGATGATGTCTTACGAAATAATATAGAAAGAGATGATTTAAACCCAGAGAAACG 380  
310 TCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGATTTGGAGATGAAATAAT 369  
381 TCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGATTTGGAGATGAAATAAT 440  
370 ACAGATTTGTACACAATGCTCCAAATTTGGTGGAAATTCGGGCAAGACAAATGAGATTACA 429  
441 ACAGATTTGTACACAATGCTCCAAATTTGGTGGAAATTCGGGCAAGACAAATGAGATTACA 500  
430 CTCACAGTGTTCCTCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGA-TGG 488  
501 CTCACAGTGTTCCTCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGA-TGG 560  
489 ATTCCAGCCCAACAGAGAAAGATCTTCAAGAAAGACCCCGCAGGAGCGGACCAAGTGAAGCC 548  
561 ATTCCAGCCCAACAGAGAAAGATCTTCAAGAAAGACCCCGCAGGAGCGGACCAAGTGAAGCC 620  
549 GTGATTTATGTACATGCGCAATGGGATTCGAAGACTGTGATGATCAGCCACCTTAAAGAA 608  
621 GTGATTTATGTACATGCGCAATGGGATTCGAAGACTGTGATGATCAGCCACCTTAAAGAA 680  
609 AGAAATCCAAAGTCAGCAAGAAAGAA 636  
681 AGAAATCCAAAGTCAGCAAGAAAGAA 708

## RESULT 15

BU591073/c  
LOCUS  
DEFINITION  
AGENCOURT\_886651 NIH\_MGC\_137 Mus musculus cDNA clone IMAGE:6430926  
5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCEBU591073  
BU591073.1 GI:23242616  
EST.Mus musculus  
Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgi.nci.nih.gov/

1 (bases 1 to 864)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gerard Gradwohl (PNAS 97 P1607-1611, 2000)

cDNA Library Preparation: Catherine Lee, Endocrine Pancreas

Consortium

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNLN at:

http://image.llnl.gov

plate: 18BD11 row: a column: 07

High quality sequence stop: 560.

Location/Qualifiers

1. 864

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:6430926"

## FEATURES

source

BASE COUNT 197 a 217 c 161 g 289 t  
ORIGIN  
Query Match 56.1%; Score 606.4; DB 13; Length 864;  
Best Local Similarity 88.4%; Pred. No. 9.7e-67;  
Matches 681; Conservative 0; Mismatches 86; Indels 3; Gaps 2;  
QY 300 AGAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGATTTGGGAG 359  
Db 767 AAAAAAACCCCTTCCAGCAGCATTTCCAGAGAGCGTTAATCAATAGCCCAAGATTTGGGAG 708  
QY 360 ATGAAAAAATACAGATTGTACACAATGCTCGAATTTGGTGGAAATTCGGGCAAGACAAA 419  
Db 707 ATGAAAAAATTCAGATTGTGCCCCAGATGCTGGAATTTGG-GGAGAACCCGAGGAGACAAA 649  
QY 420 TGGAGTTACACTCACAGTGTGTTTCCCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAG 479  
Db 648 TGGAGCTGATTCACAGTGTGTTTCCAGGATCCTGCTGAAAGTGAACGAGCCTCAGACAAGT 589  
QY 480 CAAAGATGGATTCAGCCCAACAGAAAGATCTTCAAGAAAGACCCCGCAGGAGCGGACCA 539  
Db 588 CGAAGATGGATTCAGCTCAACCGGAAAGATCTTCTAGAAGACCTCGAAGACAGAGGACCA 529  
QY 540 GTGAAGCCGCTGATTTATGTACATGGCAATTTGGGATTTGAAGACTGTGATGATCAGCCAC 599  
Db 528 GTGAGAGCCGCTGACTTATGTACATGACAAACGGGATTTGACGACTGTGATGATCAACCCC 469  
QY 600 CTAAGAAAGAAATCCAGTTCAGCAAGAAAGAAACGCTCCCAAGGGCAAGCAGGAAA 659  
Db 468 CGAAGAAAGAAATCCAGTTCAGCAAGAAAGAAAGCGCTCCCAAGGGCAAGCAGGAGA 409  
QY 660 GGGAGCTTTCACCTGTGTGAGTTTGCATATAGATCCTAATGAACCTACATATCTGTTATGCA 719  
Db 408 GGGAGCATCCCTGTGCGAGTTTGCATCGATCCCAATGAGCCACCTACTGCTTGTGTA 349  
QY 720 ACCAGTGTCTTATGGGAGATGATAGATGTGACAAATGAACAGTGTCCAAATTTGAATGCT 779  
Db 348 ACCAAGTGTCTTACGGGAGATGATAGGCTGTGACAAATGAACAGTGTCCCAATTTGAATGCT 289  
QY 780 TTCACCTTTTCATGTGTTTTCACCTTACCTATAAAGGGGAAATGCTATTGCCCCAAAGT 839  
Db 288 TTCACCTTTTCATGTGTTTTCACCTTACCTATAAAGGGGAAATGCTATTGCCCCAAAGT 229  
QY 840 GCAGGGAGATAATGAGAAAAACAATGGACAAAAGTACTGAAAAGACAAAAAGGATAGAA 899  
Db 228 GTAGGGAGACAAATGAGAAAAACAATGGACAAAAGTACTGAAAAGACAAAAAGGAGAA 169  
QY 900 GATCGAGGTAGTAAGGCCATCCACATTTTAAAGGGTATTGCTCTTTTATATAATTCGT 959  
Db 168 GAGCGAGGTAGTGAAGGCCATC--GGTTTTAAAGGGTCTTCTGCTCTTTTATAGAAATGCT 111  
QY 960 TTGCTTTTCAGAAAAATGTTTTAGGGTAAATGCAAGACTATGCAATAATTTTTTAATCATTT 1019  
Db 110 TTACTTTTCAGAAAAATGTTTTAGGGTAACTGCAATAGACTATGCAATAATTTTTTAATCATTT 51  
QY 1020 AGTATTAAATGGTGTATTAAGAAAGTTGTTGTAATTTGAAAAAAGAAAAA 1069  
Db 50 AGTATTGATGGTGTATTAAGAAAGTTGTTGTAATTTGAAAAAAGAAAAA 1

Search completed: January 12, 2004, 17:22:06

Job time: 2235 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 16:38:42 ; Search time 334 Seconds  
(without alignments)  
8728.722 Million cell updates/sec

Title: US-09-513-365A-2  
Perfect score: 1080  
Sequence: 1 gggcgccggcggtgcgtg.....aaaaaaaaaaaaaaaaaaaa 1080

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq 19Jun03:*	
1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*				
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*				
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*				
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*				
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*				
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*				
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*				
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*				
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*				
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*				
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*				
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*				
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*				
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*				
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*				
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*				
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*				
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*				
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*				
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*				
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*				
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*				
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*				
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*				
25:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1080	100.0	1080	22	AAD12783
2	1046	96.9	1078	20	AZ08596
3	838.4	77.6	840	20	AZ08595
4	825.2	76.4	1153	21	AZ47473
5	822.2	76.4	1154	24	ABS70424
6	332.2	30.8	346	21	AAC28459
7	325	30.1	325	21	AZ47474
8	253.6	23.5	2817	21	AAA53790

9	237.4	22.0	911	21	AAA53792	Human P37ING1 codi
10	236	21.9	1533	22	AAH28478	Nucleotide sequenc
11	236	21.9	2897	24	ABK86977	Human inhibitor of
12	197.2	18.3	1835	21	AAA53789	Murine ingl common
13	183.8	17.0	1143	22	AAH28479	Nucleotide sequenc
14	183.8	17.0	1902	18	AAH69651	Tumour suppressor
15	183.8	17.0	1902	19	AAV62285	Partial INGI parti
16	183.8	17.0	2061	18	AAH69652	Tumour suppressor
17	183.8	17.0	2061	19	AAV62292	Human INGI full-le
18	183.4	17.0	8487	22	AAH37089	Human musculoskele
19	183.4	17.0	8487	25	ABX60077	cDNA encoding nove
20	183.4	17.0	163350	24	AAH46127	Human tumour suppr
21	183	16.9	1905	24	AAH46126	Human tumour suppr
22	182.4	16.9	2886	24	ABK86976	Human inhibitor of
23	182.2	16.9	2061	20	AAH28688	Nucleotide sequenc
24	175	16.2	825	22	AAH03733	Human cDNA clone (
25	143	13.2	742	22	AAH28480	Nucleotide sequenc
26	142.6	13.2	857	22	AAH28481	Nucleotide sequenc
27	125.8	11.6	693	24	ABQ50058	Oligonucleotide fo
28	125.8	11.6	693	24	ABQ50059	Oligonucleotide fo
29	124.6	11.5	693	24	ABQ50060	Oligonucleotide fo
30	124.6	11.5	693	24	ABQ50061	Oligonucleotide fo
31	109.2	10.1	1082	25	ABQ77429	Human CGDD cDNA 72
32	102	9.4	108	16	AAH23370	Human gene signatu
33	101.4	9.4	1465	22	AAK52101	Human polynucleoti
34	98.8	9.1	987	23	ABL15055	Drosophila melanog
35	97	9.0	693	22	ABA08933	Human IKK binding
36	97	9.0	958	22	AAF90402	Cell cycle protein
37	97	9.0	1108	22	AAF90399	Cell cycle protein
38	97	9.0	1300	24	ABZ11521	Human polynucleoti
39	95.6	8.9	678	21	AAA27423	IkappaB kinase (IK
40	95.6	8.9	699	21	AAA27422	IkappaB kinase (IK
41	91.8	8.5	1856	22	AAI60334	Human polynucleoti
42	91.6	8.5	1336	24	ABQ54665	Human ovarian anti
43	91.6	8.5	1807	22	AAH12782	cDNA encoding huma
44	91.6	8.5	1819	25	ABZ23614	Cofactor 33a (CF33
45	91.6	8.5	1819	25	ABZ23615	Cofactor 33a (CF33

ALIGNMENTS

RESULT 1  
AAD12783  
ID AAD12783 standard; DNA; 1080 BP.

AC AAD12783;

DT 23-OCT-2001 (first entry)

XX Tumour suppressor homologue protein, p33ING2 DNA.

XX Tumour suppressor; p47ING3; cell proliferation; cellular aging; p33ING2;  
XX anchrage dependence; apoptosis; tumour; cancer; gene therapy; ds.

XX Unidentified.

XX Key Location/Qualifiers  
XX CDS 68..910  
XX /\*tag= a  
XX /product= "Tumour suppressor homologue protein, p33ING2"

XX WO200159114-A2.

XX 16-AUG-2001.

XX 09-FEB-2001; 2001WO-US04425.

XX 09-FEB-2000; 2000US-0181292.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Harris CC, Nagashima M;

XX WPI: 2001-488975/53.  
DR P-PSDB; AAE06675.  
XX New tumor suppressor protein p47ING3 for the diagnosis and treatment of  
PT tumors -  
PT  
XX  
PS Disclosure; Page 78; 80pp; English.  
XX  
XX The present sequence is a DNA encoding tumour suppressor homologue  
CC protein, p33ING2 which is homologous to human tumour suppressor protein,  
CC p47ING3. The tumour suppressors of the invention are involved in the  
CC regulation of cell proliferation and in the control of cellular aging,  
CC anchorage dependence and apoptosis. The tumour suppressor protein,  
CC p47ING3, nucleic acids encoding it and antibodies against it are useful  
CC for diagnosis, prevention and treatment of tumours and cancers. The  
CC p47ING3 DNA is also used in gene therapy.  
XX  
SQ Sequence 1080 BP; 375 A; 196 C; 271 G; 238 T; 0 other;  
  
Query Match 100.0%; Score 1080; DB 22; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 8.4e-237;  
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GCGGCCGCGCGCGTGCATGTGCGCTGCTGGATGCGGAGCGCGCGCGCGGAT 60  
DB 1 GCGGCCGCGCGCGTGCATGTGCGCTGCTGGATGCGGAGCGCGCGCGCGGAT 60  
  
QY 61 CGCGAGGATGTTAGGCGAGCAGCAGCAACTGTACTCGTGGTGGCTCTCTGACCGG 120  
DB 61 CGCGAGGATGTTAGGCGAGCAGCAGCAACTGTACTCGTGGTGGCTCTCTGACCGG 120  
  
QY 121 GGAGCGGAGCGCGCTCTACCTGCTACCTGCTGCTGAGGAGCTGCTGAGTGGTGGTGGT 180  
DB 121 GGAGCGGAGCGCGCTCTACCTGCTGCTGAGGAGCTGCTGAGTGGTGGTGGTGGT 180  
  
QY 181 GCCCAGCAGCATGCAGAGAACTGCTGCTGCTGCGAGAGCTGGAACAATAATCAAGA 240  
DB 181 GCCCAGCAGCATGCAGAGAACTGCTGCTGCTGCGAGAGCTGGAACAATAATCAAGA 240  
  
QY 241 AACGTTAAAGAAATGATGATGCTTACGAAAAATATAAGAAAGAGATGATTTAAACCA 300  
DB 241 AACGTTAAAGAAATGATGATGCTTACGAAAAATATAAGAAAGAGATGATTTAAACCA 300  
  
QY 301 GAAGAAACGCTTACAGCAGCTCTCCAGAGAGCACTAATTAATAGTCAAGAAATGGGAGA 360  
DB 301 GAAGAAACGCTTACAGCAGCTCTCCAGAGAGCACTAATTAATAGTCAAGAAATGGGAGA 360  
  
QY 361 TGAAGAAATACAGATGTTTACAAATGCTCGAATTTGGTGGAAATCGGGCAAGCAAT 420  
DB 361 TGAAGAAATACAGATGTTTACAAATGCTCGAATTTGGTGGAAATCGGGCAAGCAAT 420  
  
QY 421 GGAGTTTACACTCACAGTGTTCACAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 480  
DB 421 GGAGTTTACACTCACAGTGTTCACAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 480  
  
QY 481 AAGATGGAATTCAGCCAAACCAAGAGATCTTCAAGAGAGACCCCGAGCGAGCCAG 540  
DB 481 AAGATGGAATTCAGCCAAACCAAGAGATCTTCAAGAGAGACCCCGAGCGAGCCAG 540  
  
QY 541 TGAAGCCGCTGATTTATGTCATGCGCAATGCGATTGGATTGAGACTGTGATGATCAGCCACC 600  
DB 541 TGAAGCCGCTGATTTATGTCATGCGCAATGCGATTGGATTGAGACTGTGATGATCAGCCACC 600  
  
QY 601 TAAAGAAAGAAATCCAAGTCAGCAAGAAAGAAACGCTCCAGGCCAAGCAGGAAAG 660  
DB 601 TAAAGAAAGAAATCCAAGTCAGCAAGAAAGAAACGCTCCAGGCCAAGCAGGAAAG 660  
  
QY 661 GGAAGCTTCCACCTGTTGAGTTGCAATAGATCTTAATGAACCTACATCTGCTTATGCAA 720  
DB 661 GGAAGCTTCCACCTGTTGAGTTGCAATAGATCTTAATGAACCTACATCTGCTTATGCAA 720  
  
QY 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAATGAACAGTGTCCAATTTGAATGGTT 780

DB 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAATGAACAGTGTCCAATTTGAATGGTT 780  
QY 781 TCACCTTTTCATGCTGTTTACCTATATAACCAAGGGAATGCTATTGCCCAAGTG 840  
DB 781 TCACCTTTTCATGCTGTTTACCTATATAACCAAGGGAATGCTATTGCCCAAGTG 840  
QY 841 CAGGGGAGATATGAGAAACCAATGCAAAAGTACTGAAAAAGACAAAAAGGATAGAAG 900  
DB 841 CAGGGGAGATATGAGAAACCAATGCAAAAGTACTGAAAAAGACAAAAAGGATAGAAG 900  
QY 901 ATCGAGGTAGTAAGGCCCATCCACATTTTAAAGGGTATTGCTTTTATATAATTCGTT 960  
DB 901 ATCGAGGTAGTAAGGCCCATCCACATTTTAAAGGGTATTGCTTTTATATAATTCGTT 960  
QY 961 TCGTTTTCAGAAAATGTTTTAGGTTAAATGTCATAGACTATGCAATTAATTTTATCATTA 1020  
DB 961 TCGTTTTCAGAAAATGTTTTAGGTTAAATGTCATAGACTATGCAATTAATTTTATCATTA 1020  
QY 1021 GTATTAAATGGTGTATTAAAGTTGTTGTTACTTTGAAAAAATAAAAAAATAAAAAA 1080  
DB 1021 GTATTAAATGGTGTATTAAAGTTGTTGTTACTTTGAAAAAATAAAAAAATAAAAAA 1080  
  
RESULT 2  
AAZ08596  
ID AAZ08596 standard; cDNA; 1078 BP.  
XX  
AC AAZ08596;  
XX  
DT 18-OCT-1999 (first entry)  
XX  
DE Human INGIL encoding cDNA with 5' and 3' untranslated regions.  
XX  
KW Human; TSC403; INGIL; diagnosis; lung cancer; cell cycle; regulation;  
KW cell proliferation; cell aging; apoptosis; tumour suppressor; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 92..934  
FT /\*tag= a  
FT /product= "INGIL"  
XX  
PN WO940190-A1.  
XX  
PD 12-AUG-1999.  
XX  
PF 02-FEB-1999; 99WO-JP00419.  
XX  
PR 28-APR-1998; 98JP-0134679.  
PR 03-FEB-1998; 98JP-0038133.  
PR 05-MAR-1998; 98JP-0073234.  
XX  
PA (SAKA ) OTSUKA PHARM CO LTD.  
XX  
PI Horie M, Nagata M, Ozaki K, Shimada Y;  
XX  
DR WPI; 1999-494294/41.  
DR P-PSDB; AAY29606.  
XX  
PT Human lung-specific gene TSC430 overexpressed in cancer tissue, used  
PT for treatment of, e.g. colon tumour  
XX  
PS Claim 15; Page 91-93; 99pp; Japanese.  
XX  
CC The present sequence represents the human tumour suppressor gene INGIL.  
CC The present invention also describes the human gene TSC403 expressed  
CC specifically in normal lung tissue. TSC403 is useful in the  
CC diagnosis, investigation and treatment of cancers in which it is  
CC overexpressed, including cancer of the lung, breast, fallopian tube,  
CC oesophagus, colon, thyroid, parotid gland, bladder, ovary or pancreas.  
CC INGIL is useful in the investigation of cell proliferation, aging and

CC apoptosis and the pathology of cancer, the diagnosis and treatment of  
CC cancer such as cancer of the colon, stomach, oesophagus or fallopian  
CC tube, and the screening of candidate drugs for the treatment of such  
CC cancers.  
XX  
SQ Sequence 1078 BP; 356 A; 206 C; 280 G; 236 T; 0 other;

Query Match 96.9%; Score 1046; DB 20; Length 1078;  
Best Local Similarity 99.5%; Pred. No. 4.9e-229;  
Matches 1049; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGGCGCGCGCGTGCATGTCGGCTGCTGCGATGCGGAGCGCGCGCGCGCGGAT 60  
DB |||||  
QY 25 GCGGCGCGCGCGTGCATGTCGGCTGCTGCGATGCGGAGCGCGCGCGCGCGGAT 84  
DB |||||  
QY 61 GCGCAGGATTTAGGCGCAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTGACCGG 120  
DB |||||  
QY 85 GCGCAGGATTTAGGCGCAGCAGCAGCAGCAACTGTACTCGTGGCGCGCTCTGACCGG 144  
DB |||||  
QY 121 GGAGCGAGCGCGCTGCTACCTGCTACGTCGAGGACTACCTTGGTGGCGTGGATCGCT 180  
DB |||||  
QY 145 GGAGCGAGCGCGCTGCTACCTGCTACGTCGAGGACTACCTTGGTGGCGTGGATCGCT 204  
DB |||||  
QY 181 GCGCCAGCAGCATCCAGAGGACGTGCTGCTGCGAGAGCTCGACAACTATCAAGA 240  
DB |||||  
QY 205 GCGCCAGCAGCATCCAGAGGACGTGCTGCTGCGAGAGCTCGACAACTATCAAGA 264  
DB |||||  
QY 241 AAGCTTAAAGAAATTTGATGATGCTCTACGAAATATAGAAAGAGATGATTTAAACCA 300  
DB |||||  
QY 265 AAGCTTAAAGAAATTTGATGATGCTCTACGAAATATAGAAAGAGATGATTTAAACCA 324  
DB |||||  
QY 301 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 360  
DB |||||  
QY 325 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 384  
DB |||||  
QY 361 TGAAGAAATACAGATTTTACAAATGCTCGAAATTTGGTGGAAATCGGSCAAAGACAAAT 420  
DB |||||  
QY 385 TGAAGAAATACAGATTTTACAAATGCTCGAAATTTGGTGGAAATCGGSCAAAGACAAAT 444  
DB |||||  
QY 421 GGAAGTTACACTACAGTGTTCACAGATCTGCTGAAAGTGAACGAGCTTCAGATAAGC 480  
DB |||||  
QY 445 GGAAGTTACACTACAGTGTTCACAGATCTGCTGAAAGTGAACGAGCTTCAGATAAGC 504  
DB |||||  
QY 481 AAGATGGATTTCCAGCCAAACAGAAAGATCTTCAAGAAAGCCCGCAGCGGCGGACCCAG 540  
DB |||||  
QY 505 AAGATGGATTTCCAGCCAAACAGAAAGATCTTCAAGAAAGCCCGCAGCGGCGGACCCAG 564  
DB |||||  
QY 541 TGAAGCCGATTTATGTCATATGCAATGGCAATGGATTTGAAGACTGTGATGATCAGCCACC 600  
DB |||||  
QY 565 TGAAGCCGATTTATGTCATATGCAATGGCAATGGATTTGAAGACTGTGATGATCAGCCACC 624  
DB |||||  
QY 601 TAAAGAAAGAAATCCAGTCCAGCAAGAAAGAAACGCTTCAAGCCGAGCAGGAAAG 660  
DB |||||  
QY 625 TAAAGAAAGAAATCCAGTCCAGCAAGAAAGAAACGCTTCAAGCCGAGCAGGAAAG 684  
DB |||||  
QY 661 GGAAGCTTACCTGTTGAGTTTGCATATAGATCTTCAATGAACTTACATCTGCTTATGCAA 720  
DB |||||  
QY 685 GGAAGCTTACCTGTTGAGTTTGCATATAGATCTTCAATGAACTTACATCTGCTTATGCAA 744  
DB |||||  
QY 721 CCAAGTGTCTTATGGGAGATGATAGATGTGCAATGAACAGTGTCCAATTTGAATGGTT 780  
DB |||||  
QY 745 CCAAGTGTCTTATGGGAGATGATAGATGTGCAATGAACAGTGTCCAATTTGAATGGTT 804  
DB |||||  
QY 781 TCACCTTTTCATGTTTTCATCTACCTTATTAACCAAGGGGAAATGGTATTTGCCAAAGTG 840  
DB |||||  
QY 805 TCACCTTTTCATGTTTTCATCTACCTTATTAACCAAGGGGAAATGGTATTTGCCAAAGTG 864  
DB |||||  
QY 841 CAGGGAGATTAATGAGAAACAAATGACAAAGCTTCTGAAAGCAACAAAGGATAGAG 900  
DB |||||  
QY 865 CAGGGAGATTAATGAGAAACAAATGACAAAGCTTCTGAAAGCAACAAAGGATAGAG 924  
DB |||||  
QY 901 ATCAGGATAGTAAGGCCATCCACATTTTAAAGGGTTATTTGCTCTTTATATATTCGTT 960  
DB |||||

DB 925 ATCAGGATAGTAAGGCCATCCACATTTTAAAGGGTTATTTGACTATTTATATATCCGTT 984  
QY 961 TGCTTTTCAGAAAATGTTTGGGTAAATGCATAAGACTATGCAATATTTTATCATTA 1020  
DB 985 TGCTTTTCAGAAAATGTTTGGGTAAATGCATAAGACTATGCAATATTTATCATTA 1044

QY 1021 GTATTAAATGCTATTTAAAGTTGTTGTACTTTG 1054  
DB 1045 GTATTAAATGCTATTTAAAGTTGTTGTACTTTG 1078

RESULT 3  
AAZ08595  
ID AAZ08595 standard; cDNA; 840 BP.  
XX  
AC AAZ08595;  
XX  
DT 18-OCT-1999 (first entry)  
XX  
DE Human ING1L encoding cDNA.  
XX  
KW Human; TSC403; ING1L; diagnosis; lung cancer; cell cycle; regulation;  
KW cell proliferation; cell aging; apoptosis; tumour suppressor; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..840  
FT /\*tag= a  
FT /product= "ING1L"  
FT /note= "no stop codon given"

PN WO9940190-A1.  
XX  
PD 12-AUG-1999.  
XX  
PF 02-FEB-1999; 99WO-JP00419.  
XX  
PR 28-APR-1998; 98JP-0134679.  
PR 03-FEB-1998; 98JP-0038133.  
PR 05-MAR-1998; 98JP-0073234.  
XX  
PA (SAKA ) OTSUKA PHARM CO LTD.  
XX  
PI Horie M, Nagata M, Ozaki K, Shimada Y;  
XX  
DR WPI; 1999-494294/41.  
XX  
DR P-PSDB; AAY29606.

PT Human lung-specific gene TSC430 overexpressed in cancer tissue, used  
PT for treatment of, e.g. colon tumour  
XX  
PS Claim 14; Page 91; 99pp; Japanese.

CC The present sequence represents the human tumour suppressor gene ING1L.  
CC The present invention also describes the human gene TSC403 expressed  
CC specifically in normal lung tissue. TSC403 is useful in the  
CC diagnosis, investigation and treatment of cancers in which it is  
CC overexpressed, including cancer of the lung, breast, fallopian tube,  
CC oesophagus, colon, thyroid, parotid gland, bladder, ovary or pancreas.  
CC ING1L is useful in the investigation of cell proliferation, aging and  
CC apoptosis and the pathology of cancer, the diagnosis and treatment of  
CC cancer such as cancer of the colon, stomach, oesophagus or fallopian  
CC tube, and the screening of candidate drugs for the treatment of such  
CC cancers.

SQ Sequence 840 BP; 297 A; 164 C; 211 G; 168 T; 0 other;

Query Match 77.6%; Score 838.4; DB 20; Length 840;  
Best Local Similarity 99.9%; Pred. No. 1.1e-181;  
Matches 839; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 68 ATGTTAGGCGCAGCAGCAGCAACTGTACTCGTCGGCTCGCTCTGACCGGGAGCGG 127

```

Db      1  ATGTTAGGCGAGCAGCAGCAGCAACTGTAAGTCTGTCGCGCGCTCTCTGCGCGGAGCGG 60
QY      128 AGCGGCTGCTACCTGCTACGTCGAGGACTACCTGAGTCGTCGAGTCGTCGCGCCAC 187
Db      61  AGCGGCTGCTACCTGCTACGTCGAGGACTACCTGAGTCGTCGAGTCGTCGCGCCAC 120
QY      188 GACATCGAGGAACTGTCGTCGTCGCGAGAGCTGACCAACAATATCAAGAAAGCTTA 247
Db      121 GACATCGAGGAACTGTCGTCGTCGCGAGAGCTGACCAACAATATCAAGAAAGCTTA 180
QY      248 AAGGAAATGATGATGTCACGAAATAATAAGAAAGAGATGATTTAAACCAAGAGAAA 307
Db      181 AAGGAAATGATGATGTCACGAAATAATAAGAAAGAGATGATTTAAACCAAGAGAAA 240
QY      308 CGTCTACAGAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGATGAAA 367
Db      241 CGTCTACAGAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGATGAAA 300
QY      368 ATACAGATTTTACACAAATGCTCGAATTTGGTGGAAATCGGCAAGACAAATGGAGTTA 427
Db      301 ATACAGATTTTACACAAATGCTCGAATTTGGTGGAAATCGGCAAGACAAATGGAGTTA 360
QY      428 CACTCAGATTTTCAAGATCTGCTGAAAGTGAACGAGCTCAGATAAAGCAAGATG 487
Db      361 CACTCAGATTTTCAAGATCTGCTGAAAGTGAACGAGCTCAGATAAAGCAAGATG 420
QY      488 GATTCCAGCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGGCGGACCAAGTGAAGC 547
Db      421 GATTCCAGCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGGCGGACCAAGTGAAGC 480
QY      548 CGTGATTTATGTCACATGCAATGGATTTGAAGCTGTGATGATCAGCCACCTAAGAA 607
Db      481 CGTGATTTATGTCACATGCAATGGATTTGAAGCTGTGATGATCAGCCACCTAAGAA 540
QY      608 AAGAAATCCAAGTCAGCAAGAAAGAAACGCTCCAAGGCCAAGCAGGAAAGGGAAGCT 667
Db      541 AAGAAATCCAAGTCAGCAAGAAAGAAACGCTCCAAGGCCAAGCAGGAAAGGGAAGCT 600
QY      668 TCACCTGTTGAGTTTCAATAGATCTTAATGAACCTACATGCTTATGCAACCAAGTG 727
Db      601 TCACCTGTTGAGTTTCAATAGATCTTAATGAACCTACATGCTTATGCAACCAAGTG 660
QY      728 TCTTATGGGAGATGATAGATGTGCAATGAACAGTGTCCATTTCAATGGTTTCACTTT 787
Db      661 TCTTATGGGAGATGATAGATGTGCAATGAACAGTGTCCATTTCAATGGTTTCACTTT 720
QY      788 TCATGTTTCACTTACCTATAACCAAGGGGAAATGGTATTGCCAAAGTCAGGGGA 847
Db      721 TCATGTTTCACTTACCTATAACCAAGGGGAAATGGTATTGCCAAAGTCAGGGGA 780
QY      848 GATAATGAGAAACAATGGCAAAAGTACTGAAAAGACAAAAGGATGAAGATCGAGG 907
Db      781 GATAATGAGAAACAATGGCAAAAGTACTGAAAAGACAAAAGGATGAAGATCGAGG 840

```

## RESULT 4

AA247473  
ID AA247473 standard; cdna; 1153 BP.

AC AA247473;

XX AA247473;

DE 14-MAR-2000 (first entry)

XX Human tumour suppressor (TUSUP) nucleotide sequence.

XX Tumour suppressor protein; TUSUP; human; cancer; treat; prevent;

XX reproductive tract; gastrointestinal tract; immune system; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 277..999

FT /\*tag= a  
FT /product= TUSUP  
FT /note= "Tumour suppressor protein"

XX WO9961612-A1.

XX 02-DEC-1999.

XX 20-MAY-1999; 99WO-US11136.

XX 28-MAY-1998; 98US-0086359.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Patterson C;

XX WPI; 2000-062711/05.

DR P-PSDB; AAY52199.

XX New human tumour suppressor protein for treating cancer, particularly of reproductive and gastrointestinal tracts or immune system

PS Claim 7; Fig 1; 71pp; English.

XX This sequence encodes the human tumour suppressor (TUSUP) protein. The TUSUP sequence was identified from a human lung cDNA library. The protein is 240 amino acids in length and has one potential CAMP and cGMP dependent protein kinase phosphorylation site, nine potential protein kinase C phosphorylation sites, five potential tyrosine kinase phosphorylation sites, and one potential tyrosine kinase phosphorylation site. TUSUP has structural and chemical similarity to p33. TUSUP can be used to treat or prevent cancer, particularly of the reproductive or gastrointestinal tracts or of the immune system. Agonists of the TUSUP protein can be used to raise TUSUP specific antibodies and to screen for specific binding agents and potential therapeutic agents. The TUSUP nucleotide sequence can be used in gene therapy and as a source of diagnostic probes and primers.

XX Sequence 1153 BP; 367 A; 215 C; 285 G; 265 T; 21 other;

Query Match 76.4%; Score 825.2; DB 21; Length 1153;

Best Local Similarity 97.9%; Pred. No. 1.3e-178;

Matches 836; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 211 GCTCGAGAGCTGGCAACAATAATATCAAGAAACGTTAAAGGAATGTGATGCTCTACGA 270

Db 300 GCTCGAGAGCTGGCAACAATAATATCAAGAAACGTTAAAGGAATGTGATGCTCTACGA 359

QY 271 AAAATATATAAGAAAGAGATGATTTAAACCAAGAAACGCTTACAGCAGCTTCTCCAGAG 330

Db 360 AAAATATATAAGAAAGAGATGATTTAAACCAAGAAACGCTTACAGCAGCTTCTCCAGAG 419

QY 331 AGCAGCTAATTAATAGTCAAGAAATGGGAGATGAAAAATACAGATTGTTTACAAATGCT 390

Db 420 AGCAGCTAATTAATAGTCAAGAAATGGGAGATGAAAAATACAGATTGTTTACAAATGCT 479

QY 391 CGAATTTGGTGAAGAAATCGGGCAAGCAAAATGAGTTACACTCACAGTGTTCAGAGATCC 450

Db 480 CGAATTTGGTGAAGAAATCGGGCAAGCAAAATGAGTTACACTCACAGTGTTCAGAGATCC 539

QY 451 TCGTGAAGTGAACGAGCTCAGATAAAGCAAGATGGATTCCAGCAACCAAGAGATC 510

Db 540 TCGTGAAGTGAACGAGCTCAGATAAAGCAAGATGGATTCCAGCAACCAAGAGATC 599

QY 511 TTCAAGAAAGACCCCGCAGGACCGGACCAAGTGAAGCCGCTGATTATGTACATGGCAAA 570

Db 600 TTCAAGAAAGACCCCGCAGGACCGGACCAAGTGAAGCCGCTGATTATGTACATGGCAAA 659

QY 571 TGGGATTTGAAGACTGTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAGTCAGCAAGAA 630

Db 660 TGGGATTTGAAGACTGTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAGTCAGCAAGAA 719

QY 631 AAAGAAACGCTCCAAAGGCCAAGCAAGGGAAGCTTCCACCTGTTGAGTTTGCATAGAGA 690

Db 720 AAAGAAACGCTCAAGGCCAAGCAGGAAAGGGAGCTTCACCTGTTGAGTTTGCATAGA 779  
QY TCCTAATGAACCTACATACCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATG 750  
Db 780 TCCTAATGAACCTACATACCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATG 839  
QY 751 TGCAATGAACAGTGTCCAAATGAATGGTTTCACTTTTTCATGTTTCACTTACCTATAA 810  
Db 840 TGCAATGAACAGTGTCCAAATGAATGGTTTCACTTTTTCATGTTTCACTTACCTATAA 899  
QY 811 ACCAAGGGGAATGTTATGCCCCAAGTGCAGGGGAGATTAATGAGAAAACAATGACAA 870  
Db 900 ACCAAGGGGAATGTTATGCCCCAAGTGCAGGGGAGATTAATGAGAAAACAATGACAA 959  
QY 871 AAGTACTGAAAAGACAAAAGGATAGAGATGAGTGTAGTAAAGGCCATCCACATTTTA 930  
Db 960 AAGTACTGAAAAGACAAAAGGATAGAGATGAGTGTAGTAAAGGCCATCCACATTTTA 1019  
QY 931 AAGGGTTATTGTTCTTTTATATATGTTTGTCTTTTTCAGAAAATGTTTATAGGGTAAATGC 990  
Db 1020 AAGGGTTATTGTTCTTTTATATATGTTTGTCTTTTTCAGAAAATGTTTATAGGGTAAATGC 1079  
QY 991 ATAAGACTATGCAATATTTTAAATCATATAGTATTAATGTTGTTTAAAGTTGTTGTAC 1050  
Db 1080 ATAAGACTATGCAATATTTTAAATCATATAGTATTAATGTTGTTTAAAGTTGTTGTAC 1139  
QY 1051 TTTGAAAAAATA 1064  
Db 1140 TTTGAAAAAATA 1153

RESULT 5

ABST0424  
ID ABST0424 standard; cdna; 1154 BP.

XX AC ABST0424;  
XX DT 27-NOV-2002 (first entry)  
XX DE Human bone remodelling gene #81.  
XX KW Bone remodelling; osteoporosis; human; gene; ss.

XX OS Homo sapiens.  
XX PN US6426186-B1.

XX PD 30-JUL-2002.

XX PF 18-JAN-2000; 2000US-0484970.

XX PR 18-JAN-2000; 2000US-0484970.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Jones KA, Volkmut W, Walker MG;

XX PF WPI; 2002-673014/72.

XX A combination of polynucleotides which are co-expressed with genes  
PT known to be involved in bone remodeling and osteoporosis are useful in  
PT an array for the diagnosis of bone remodeling and osteoporosis  
PT associated disorders -

XX Claim 1; Column 239-242; 206pp; English.

XX The invention relates to a combination comprising a number of  
CC substantially purified and isolated polynucleotides which are  
CC co-expressed with genes known to be involved in bone remodeling and  
CC osteoporosis. The invention is used to diagnose disorders associated  
CC with bone remodeling or osteoporosis. ABST0344-ABS70512 represent  
CC human bone remodelling genes of the invention.

XX SQ Sequence 1154 BP; 367 A; 216 C; 285 G; 265 T; 21 other;  
Query Match 76.4%; Score 825.2; DB 24; Length 1154;  
Best Local Similarity 97.9%; Pred. No. 1.3e-178;  
Matches 836; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
QY 211 GCTCGAGAGCTGGCAACAATAATCAAGAAACGTTTAAAGGAATTTGATGATCTACGA 270  
Db 301 GCTCGAGAGCTGGCAACAATAATCAAGAAACGTTTAAAGGAATTTGATGATCTACGA 360  
QY 271 AAAATATAAGAAAGAGATGATTTAAACCAAGAAACGTTTAAAGGAATTTGATGATCTACGA 330  
Db 361 AAAATATAAGAAAGAGATGATTTAAACCAAGAAACGTTTAAAGGAATTTGATGATCTACGA 420  
QY 331 AGCAGCTAAATTAATAGTCAAGAAATTTGGGAGATGAAAAATACAGATTTGTACAAATGCT 390  
Db 421 AGCAGCTAAATTAATAGTCAAGAAATTTGGGAGATGAAAAATACAGATTTGTACAAATGCT 480  
QY 391 CGAATTTGGTGGAAAAATCGGGCAAGACAAAATGGAAGTTTACACTCAAGTGTGTTTCAAGATCC 450  
Db 481 CGAATTTGGTGGAAAAATCGGGCAAGACAAAATGGAAGTTTACACTCAAGTGTGTTTCAAGATCC 540  
QY 451 TGTGTAAGTGAACGAGCCTCAGATTAAGCAAGATGGAATTCAGGCCAACCAAGATGATC 510  
Db 541 TGTGTAAGTGAACGAGCCTCAGATTAAGCAAGATGGAATTCAGGCCAACCAAGATGATC 600  
QY 511 TTCAGAAAGACCCCGCAGGACCGGACCAAGTGAAGCCGTTTATGTCACATGCAAA 570  
Db 601 TTCAGAAAGACCCCGCAGGACCGGACCAAGTGAAGCCGTTTATGTCACATGCAAA 660  
QY 571 TGGGATTGAAGACTGTGATGATCAGCCACCTTAAAGAAAGAAATCCAAGTCAGCAAGAA 630  
Db 661 TGGGATTGAAGACTGTGATGATCAGCCACCTTAAAGAAAGAAATCCAAGTCAGCAAGAA 720  
QY 631 AAAGAAACGCTCCAGGCCAAGCAGGAAAGGGAGCTTCACCTGTTGAGTTTGCATATAGA 690  
Db 721 AAAGAAACGCTCCAGGCCAAGCAGGAAAGGGAGCTTCACCTGTTGAGTTTGCATATAGA 780  
QY 691 TCCTAATGAACCTACATACCTGTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATG 750  
Db 781 TCCTAATGAACCTACATACCTGTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATG 840  
QY 751 TGCAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCACTGTTTCACTTACCTATAA 810  
Db 841 TGCAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCACTGTTTCACTTACCTATAA 900  
QY 811 ACCAAAGGGGAAATGTTATGCCCCAAGTGCAGGGGAGATTAATGAGAAAACAATGACAA 870  
Db 901 ACCAAAGGGGAAATGTTATGCCCCAAGTGCAGGGGAGATTAATGAGAAAACAATGACAA 960  
QY 871 AAGTACTGAAAAGACAAAAGGATAGAGATGAGTGTAGTAAAGGCCATCCACATTTTA 930  
Db 961 AAGTACTGAAAAGACAAAAGGATAGAGATGAGTGTAGTAAAGGCCATCCACATTTTA 1020  
QY 931 AAGGGTTATTGTTCTTTTATATATGTTTGTCTTTTTCAGAAAATGTTTATAGGGTAAATGC 990  
Db 1021 AAGGGTTATTGTTCTTTTATATATGTTTGTCTTTTTCAGAAAATGTTTATAGGGTAAATGC 1080  
QY 991 ATAAGACTATGCAATATTTTAAATCATATAGTATTAATGTTGTTTAAAGTTGTTGTAC 1050  
Db 1081 ATAAGACTATGCAATATTTTAAATCATATAGTATTAATGTTGTTTAAAGTTGTTGTAC 1140  
QY 1051 TTTGAAAAAATA 1064  
Db 1141 TTTGAAAAAATA 1154

RESULT 6

AAC28459  
ID AAC28459 standard; cdna; 346 BP.  
XX AAC28459  
AC AAC28459;





QY 970 AAAATGTTTGGGTAATGCAATGACTATGCAATTAATTTTAATCATTAGTATTAATG 1029  
Db 85 AAAATGTTTGGGTAATGCAATGACTATGCAATTAATTTTAATCATTAGTATTAATG 26  
QY 1030 GTGTATTAAAGTGTGTACTTTG 1054  
Db 25 GTGTATTAAAGTGTGTACTTTG 1

RESULT 8

AAA53790

ID AAA53790 standard; cDNA; 2817 BP.

XX AAA53790;

XX 19-DEC-2000 (first entry)

XX Murine P37ING1 coding sequence.

DE p53; tumour; cancer; detection; antibody; hybridisation; exon 1b;  
KW INGI1; ing1; p37ING1; p37ING1; oncogene; gene therapy; diagnosis;  
KW proliferation disorder; transformation; transformed cell; mouse; ds.  
XX Mus musculus.

XX Key Location/Qualifiers  
FH 847..1686  
FT /\*tag= a  
FT /product= P37ING1 polypeptide

XX W0200046370-A1.

XX 10-AUG-2000.

XX 04-FEB-2000; 2000WO-US02959.

XX 04-FEB-1999; 99US-0118941.

XX (UNII ) UNIV ILLINOIS FOUND.

XX Gudkov A, Zeremski M, Gurova KV, Grigorian IA;

XX WPI; 2000-491278/43.

XX P-PSDB; AAY97242.

XX Detecting nucleic acid encoding exon 1b of ing1, useful for diagnosing  
PT and creating cancer, comprises contacting sample with isolated nucleic  
PT acid comprising sequence of exon 1b and detecting hybridized products  
XX Disclosure; Fig 11; 134pp; English.

XX Mutations in or loss of the p53 gene occur in more than 50% of  
CC human tumours and tumour cell lines, but functional inactivation of  
CC the p53 pathway occurs in a much larger proportion of tumours. In  
CC many cases the mechanism of functional inactivation of the p53 gene  
CC remains unknown but p53 has been found to act in cooperation with  
CC INGI1. Functional cooperation between INGI1 and p53 suggested that  
CC INGI1 encoded a tumour suppressor protein that functioned within the  
CC p53 pathway. This data suggested a possible role for INGI1 in head  
CC and neck cancers and chromosomal location of the INGI1 placed it  
CC within a region that is frequently rearranged in head and neck  
CC cancers. Large scale analysis of tumours involving INGI1 has not  
CC revealed mutations in INGI1 nor significant variations in its  
CC expression suggesting that INGI1 was not a useful gene to study in  
CC cancer etiology. However, alternative initiation exons of the ing1  
CC gene, each having their own promoter have been discovered.  
CC Expression of one promoter (1a) produces a protein identical to  
CC INGI1. Expression of a second promoter (1b) produces a protein having  
CC an identical C-terminal fragment to INGI1 but an additional 104  
CC N-terminal amino acids. The newly discovered protein has been  
CC designated p37ING1 (Wild type: p37ING1). p37ING1 has the  
CC characteristics of an oncogene. When overexpressed in cells (even  
CC those expressing wild type p53) p37ING1 is able to cause

CC proliferation or transformation of those cells. Thus detecting a  
CC nucleic acid encoding exon 1b of ing1 by hybridisation with an  
CC isolated nucleic acid having the sequence of exon 1b of ing1  
CC or its antisense sequence can identify individuals expressing the  
CC oncogenic form of ing1. Novel peptide sequences taken from the 104  
CC N-terminal peptide of p37ING1 can also be used to raise antibodies  
CC that can also be used in detection methods for the p37ING1 variant.  
CC The polypeptides may be useful in gene therapy for treatment of cell  
CC proliferation disorders, especially cancers and for diagnosing and  
CC studying cancers.

XX  
SQ Sequence 2817 BP; 670 A; 711 C; 846 G; 590 T; 0 other;

Query Match 23.5%; Score 253.6; DB 21; Length 2817;  
Best Local Similarity 60.3%; Pred. No. 5.8e-48;  
Matches 490; Conservative 0; Mismatches 289; Indels 33; Gaps 3;

QY 116 ACCGGGAGCGGAGCCGGTCTCCTCAGTCTGAGGAGTACCTTTGAGTGGTGGAG 175  
Db 862 AACGGGAGCAGATCCACCTGGTGA---ACTATGTGGAGGATTACCTGGAGTCAATCGAG 918  
QY 176 TCGCTCCCCCAGACATGAGAGGACGCTGCTGCTGCGAGAGCTGGACACAATATAT 235  
Db 919 TCACTGCCTTTTCACCTGCGAGAGGAACTCTCGCTGATGCGGAGATCGACGCCAATATC 978  
QY 236 CAAGAAACGTTAAAGGAAATTTGATGATGCTTACGAAAAATATAGAAAGAGATGATTTA 295  
Db 979 CAAGAGATCTTGAAGAGCTGGAGACTACTATGAGAGTTCAACGGGAGACAGACGGC 1038  
QY 296 AACCAAGAAACGTTCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGATTG 355  
Db 1039 ACCCAAGCGCGGGTACTGCACTGATCCAGAGGCGCTGATCCGACCCAGGAGCTA 1098  
QY 356 GAGATGAAAGAAATACAGATTGTTACAAATGCTCGAATTTGTGGAATTCGGGCAAGA 415  
Db 1099 GCGATGGAAGATCCAGATCGTGAGTCAAGTGGTGGAGCTGGTGGAGAACCCGACGAGA 1158  
QY 416 CAAATGGAGTTACACTCACAGTGTGTTTCCAAAGATCTCTGCTGAAA----- 458  
Db 1159 CAGTGGACAGTCACTGAGGAGCTTTTGAAGACACACAGGACATCATGACGGCAGCTGGT 1218  
QY 459 -----GTGAACGAGCCTCAGATAAAGCAAGAGATGGATTCCAGCCCAACAGAAAGATCT 511  
Db 1219 GGCAGCGCAAGCGCGGCGGAGCAAGTCAAGAGTGAGGSCCATCACACAGGAGGAGTAAG 1278  
QY 512 TCAGAA-----GACCCCGAGGAGCGGACAGTGAAGCCGCTGATTATGTCAATG 565  
Db 1279 CCGAATAACAAGCGGTCCAGAGGAGCGCAAAACAATGAGAAATCGAGAGAACGGGTGCAAT 1338  
QY 566 GCAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAGTCAAGCA 625  
Db 1339 AATCAGACCATGATGACATCACTCAGGAACCCCAAGGAGAGAAAGAAACCACTCA 1398  
QY 626 AAGAAAGAAAGAGCTCCAGGCGCAAGAGGAGGAGGAGGAGCTTCACTGTGTGAGTTTGA 685  
Db 1399 AAGAAGAAAGAGCTCCAGGCGCAAGGAGGAGGAGGAGGAGCTCTCTGCGGACCTTCCC 1458  
QY 686 ATAGATCTTAATGAACCTTACATCTCTTATGCAACAAGTGTCTTATGGGAGAGATGA 745  
Db 1459 ATGACCCCAAGAGCGCCACGCTACTCTGTGTCAACACGAGTCTCTTACGGGAGAGATG 1518  
QY 746 GAGTGTGACAAATGAACAGTGTCCAATGATGTTGTTTCACTTTTCATGTTTTCATTACC 805  
Db 1519 GCGTGTGACAAACAGCAATGCCCCCATCGAGTGGTTTCCACTTCTTCCGCGGGGCTCAAC 1578  
QY 806 TATAAACCAAGGGGAAATGGTATTGCCCAAGTCAGGGGAGATATAGAGAAAAAATG 865  
Db 1579 CATTAACCAAGGGGCAAGTGGTACTGCCCCCAAGTGGCGTGGGAGAGCGGAGAACCATG 1638  
QY 866 GACAAAAGTACTGAAAGACAAAAAGAGATAG 897  
Db 1639 GACAAAGCCCTGGAGAAATCCCAAGAAAGAGAG 1670





```

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 451..1290
FT /tag= a
FT /transl_except= "(pos: 817..819, aa: Val)"
FT /transl_except= "(pos: 835..837, aa: Val)"
FT /transl_except= "(pos: 853..855, aa: Asn)"
FT /transl_except= "(pos: 859..861, aa: Asp)"
FT /transl_except= "(pos: 865..867, aa: Val)"
FT /transl_except= "(pos: 874..876, aa: Ser)"
FT /product= "cancer associated antigen"
XX
PN W0200147959-A2.
XX
XX 05-JUL-2001.
XX
XX 29-NOV-2000; 2000WO-US42334.
XX
XX 30-NOV-1999; 99US-0451739.
XX 24-OCT-2000; 2000US-0602362.
XX
XX (LJDM-) LUDWIG INST CANCER RES.
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;
XX WPI; 2001-441706/47.
XX P-PSDB; AAB84698.
XX
XX Isolated cancer associated nucleic acid molecule identified by SEREX
XX (serological identification of antigens by recombinant expression
XX cloning) technique, useful in nucleic acid based therapies to treat
XX cancer -
XX
XX Claim 1; Page 43-44; 62pp; English.
XX
XX The present sequence encodes a human cancer associated antigen.
XX The sequence is a variant of the ING1 gene, which is a tumour
XX suppressor gene candidate. The cancer associated antigen polynucleotides
XX and polypeptides are useful for screening for the possible presence of
XX a pathological condition in a subject such as cancer. The cancer
XX associated antigen polypeptides are useful for producing vaccines.
XX
XX Sequence 1533 BP; 336 A; 431 C; 521 G; 244 T; 1 other;
XX
XX Query Match 21.9%; Score 236; DB 22; Length 1533;
XX Best Local Similarity 58.9%; Pred. No. 4.8e-44;
XX Matches 462; Conservative 0; Mismatches 290; Indels 32; Gaps 2;
XX
XX 145 CTAGCTGCAGGACTACCTTGAGTCGCTGGAGTCGCTGCCCGCAGCATGCAGAGGAACTG 204
XX 492 CTATGTGGAGGACTACTCGACTCCATCGAGTCCTCCCTTCGCTTCGAGAGAAATGT 551
XX
XX 205 GTCTGTGCTCGAGAGCTGGACAACTATCAAGAAACCTTAAAGGAAATTCATGATGT 264
XX
XX 552 CTGCTGTATCGGGAGATCGACCGGAATACCAAGAGATCCTGAAGAGCTAGACGATG 611
XX
XX 265 CTACGAAAAATATAAGAAAGAGATGATTTAAACCCAGAGAAACGCTTACAGCAGCTTCT 324
XX
XX 612 CTACGAGCGCTTCAGTCGCGAGACAGACGGCGCGCAGAGCGCGGATGCTGCACTGTGT 671
XX
XX 325 CCAGAGAGCACTAATTAATAGTCAAGATTTGGAGATGAAAAATACAGATTGTTACACA 384
XX
XX 672 GCAGCGCGGCTGATCCGACGCCAGGAGCTGGCGCAGCAGAGATCCAGATCTGTGAGCCA 731
XX
XX 385 AATGCTCGAATTCGTGAAATCGGCAACACAAATGGAGTTACACTCAGAGTTTCCA 444
XX
XX 732 GATGTGGAGCTGTTGGAGAACCGCGCGCAGGTTGGACGACGCTGGAGCTGTTCGA 791
XX
XX 445 AGATC-----CTGCTGAAAGTGAACGAGCCTCAG 473
XX

```

---

```

Db 792 GCGCAGCAGGAGCTGGGCGACACAGCGGGCAACACGCGGCAAGGCTGGCGCGACAGGCC 851
QY 474 ATAAAGCAAGATGGATTCCAGCCAAACAGAAAGATCTTCAAGAAAGACCCCGCAGGCGAGC 533
Db 852 CAAAGCGGAGGCGGCGAGCGCAGCGCTGACAAGCCCAACAGCAGCGCTCAGGCG-GGCAGC 910
QY 534 GGACCAAGTGAAGCCCGTGAATTTATGTCAATGGCAAAATGGGATTTGAAGACTGTGATGATC 593
Db 911 GCAACAAACGAGAAACCGTGAGAACCGCTCCAGCAACCCACGACCAACGCGCGCTCGG 970
QY 594 AGCCACCTAAAGAAAGAAATCCAGTCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 653
Db 971 GCACACCCAGGAGAGAGGCGCAAGCTCCAAAGAAAGAAAGAGCGCTCCAGGCGCAAGG 1030
QY 654 AGGAAGGGAAGCTTCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATATCTGCT 713
Db 1031 CGGAGCGAGAGGCGTCCCTGCGACCTCCCATCGACCCCAACGAAACCCACGCTACTGTC 1090
QY 714 TATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTCAATGAACAGTGTCTCAATTG 773
Db 1091 TGTGCAACCAAGTGTCTTATGGGAGATGATCGGTGCGACCAACGACGAGTGTGCTCTGTC 1150
QY 774 AATGTTTCACTTTTCATGTGTTTCACTTACCTATAAACCAAGGGAATGTTATGTC 833
Db 1151 AGTGGTTCCACTTCTCGTGGGCTCAATCATATAACCCAGGCGCAAGTGTGCTGTC 1210
QY 834 CAAAGTGCAGGAGAGATATGAGAAACAAATGAGACAAAGTACTGAAAAGACAAAAAGG 893
Db 1211 CCAAGTGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1270
QY 894 ATAG 897
Db 1271 AGAG 1274

```

---

```

RESULT 11
ABK86977
ID ABK86977 standard; cDNA; 2897 BP.
XX
AC ABK86977;
XX
DT 24-SEP-2002 (first entry)
XX
DE Human inhibitor of growth 1 (ING1) isoform, p33ING1b, cDNA.
XX
XX Human; gene; ss; monoclonal antibody; epitope; inhibitor of growth 1;
XX ING1; tumour suppressor; cell cycle; breast tumour; lymphoid malignancy;
XX cell growth; anti-ING1; CAB; isoform; diagnosis; tumour; antigen;
XX p33ING1b.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 873..1712
XX /*tag= a
XX /product= "p33ING1b"
XX
XX US2002039735-A1.
XX
XX 04-APR-2002.
XX
XX 04-JUN-2001; 2001US-0874347.
XX
XX 02-JUN-2000; 2000US-208829P.
XX 26-FEB-1999; 99US-0258372.
XX 15-NOV-1996; 96US-0751230.
XX 08-DEC-1995; 95US-0569721.
XX 22-MAR-2000; 2000US-0532868.
XX
XX (RIAB/) RIABOWOL K T.
XX (BOLA/) BOLAND D.
XX
XX Riabowol KT, Boland D;

```







XX AAV62285;  
 XX 18-JAN-1999 (first entry)  
 XX Partial INGI partial cDNA sequence.  
 XX INGI gene; p33ING1; human; apoptosis; cell death; breast cancer;  
 KW brain tumour; gene therapy; tumour suppressor; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT 109..741  
 FT /\*tag= a  
 XX  
 XX W09844102-A2.  
 XX  
 XX 08-OCT-1998.  
 PD  
 XX 26-MAR-1998; 98WO-CA00277.  
 XX  
 XX 27-MAR-1997; 97US-0828158.  
 XX (UYTE-) UNIV TECHNOLOGIES INT INC.  
 PA  
 XX Garkavtsev I, Helbing CC, Johnston RN, Riabowol K;  
 PI  
 XX WPI: 1998-542700/46.  
 DR P-PSDB; AAW79674.  
 DR  
 XX  
 PT Modulating eukaryotic apoptosis by increasing p33ING1 activity -  
 PT using p33ING1 derivatives, to induce apoptosis in cancer cells, and  
 PT in the investigation of apoptotic pathways  
 PT  
 XX Example 2; Fig 2; 66pp; English.  
 XX  
 CC This is the nucleotide sequence of a human INGI1 (INHIBITOR of  
 CC Growth) partial cDNA clone that codes for a p33ING1 polypeptide  
 CC (see AAW79674), a novel inhibitor of cell growth and a candidate  
 CC tumour suppressor. INGI1 is a new gene that is expressed in normal  
 CC mammary epithelial cells, but which is expressed only at lower  
 CC levels in several cancerous mammary epithelial cell lines and is  
 CC not expressed in many primary brain tumours. To isolate INGI1, a  
 CC subtractive hybridisation of breast cancer cell line cDNAs was  
 CC performed with cDNA from normal mammary epithelial cells, and  
 CC subcloned cDNAs were cloned into retrovirus vector pLNCX.  
 CC Following passage through a packaging line, normal mouse mammary  
 CC epithelial cells were infected, and infected cells were injected  
 CC into nude mice. Putative transforming fragments from tumours were  
 CC isolated by PCR (see AAV62290-91) and subcloned into LNCX. An INGI1  
 CC fragment was obtained and used to screen normal human fibroblast  
 CC and HeLa cell cDNA libraries. 2 Clones were sequenced to obtain  
 CC the partial INGI1 sequence. The complete cDNA sequence (see  
 CC AAV62292) was obtained by RACE. A claimed method to potentiate  
 CC apoptosis in a eukaryotic cell involves administering an active  
 CC p33ING1 peptide or an oligonucleotide encoding such as a peptide.  
 CC A claimed method for inhibiting apoptosis in a eukaryotic cell  
 CC involves administering an antisense oligonucleotide. Also claimed  
 CC are a method for determining the apoptotic characteristics of a  
 CC eukaryotic cell, an assay for determining the level of p33ING1  
 CC activity in a eukaryotic cell, and an isolated eukaryotic cell  
 CC substantially free of p33ING1 biological activity. The invention  
 CC discloses INGI1 derivatives or variants that may be used to induce  
 CC apoptosis in eukaryotic cancer cells.  
 XX  
 SQ Sequence 1902 BP; 574 A; 390 C; 462 G; 476 T; 0 other;

Query Match 17.0%; Score 183.8; DB 19; Length 1902;  
 Best Local Similarity 57.0%; Pred. No. 4.3e-32;  
 Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;  
 QY 237 AAGAACGTTAAAGGAAATTGATGATCTACGAAAAATATAGAAAGAGATGATTAA 296

Db 35 AACAGATCCTGAAGGAGCTAGACGAGTGTCTACGAGCGCTTCAGTCCGAGACAGACGGGG 94  
 QY 297 ACCAGAAGAAACGTCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGAAATTGG 356  
 Db 95 CCGAGAGCGCGGATCTGCACTGTGTGACGCGCGCTGTATCCGACGACGAGAGCTGG 154  
 QY 357 GAGATGAAAAAATACAGATTGTTTACACAAATGCTCGAAATTGGTGGAAAAATCGGGCAGAC 416  
 Db 155 GCGACGAGAAGATCCAGATCGTGTGAGCTGTGTGAGCTGTGTGGAGAACCGCACGCGGC 214  
 QY 417 AAATGGAGTTACTACTCAGTGTTCCTCAAGATC----- 449  
 Db 215 AGGTGACACGCCACGCTGGAGCTGTTTCGAGGCGCAGCAGGAGCTGGGCGCACACAGTGGCA 274  
 QY 450 -CTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATGATTCACAGCCAAACAGAAAGA 508  
 Db 275 ACAGCGGCMAGGTGGCGCGGACAGGCCCAATGGCGATGCGGTAGCGCATCTGTGACAGC 334  
 QY 509 TCTTCAAGAA--GACCCCGCAGCGCAGCAGTGAAGAGCCGTGATTTATGTTCATGG 566  
 Db 335 CCAACAGCAAGCGCTCACGGCGCAGCGCAACAACGAGAACCGTGAGAACCGCGTCCAGCA 394  
 QY 567 CAAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAGTCAAGCAA 626  
 Db 395 ACCACGACCACGACGCGCGCTCGGGCACACCCCAAGGAGAGAGGCGCAAGACCTCCA 454  
 QY 627 AGAAAAAGAAACGCTCCAGGCGCAGCAGGAGGAGGAGGAGCTTCACCTGTTGAGTTTGC 686  
 Db 455 AGAAGAAGAAAGCGCTCCAAAGGCCAAGGCGGAGGAGGCGCTCCCTGCCGACCTCCCA 514  
 QY 687 TAGATCCTTAATGAACCTACATACTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746  
 Db 515 TCGACCCCAACGAAACCCAGTACTGTCTGTGCAACACAGGTCTCTATGGGAGATGATCG 574  
 QY 747 GATGTGACAAATGAACAGTGTCCAAATTGAATGTTTTCACCTTTTTCATGTGTTTACCT 806  
 Db 575 GCTGGGACAAACGACGAGTGCCCATCGAGTGGTTCCACTTCTCGTGGTGGGCTCAATC 634  
 QY 807 ATAAACCAAGGGGAAATGGTATTGCCCAAGTGCAGGGGAGATATAGAAAAACATGG 866  
 Db 635 ATAAACCCCAAGGGCAAGTGGTACTGTCCCAAGTGCAGGGGAGAGACGAGAGACCATGG 694  
 QY 867 ACAAAAGTACTGAAAAAGACAAAAAGGATAG 897  
 Db 695 ACAAGCCCTGGAGAAATCCAAAAAGAGAG 725

Search completed: January 12, 2004, 16:44:42  
 Job time : 344 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 16:38:42 ; Search time 3978 Seconds  
(without alignments)  
11106.686 Million cell updates/sec

Title: US-09-513-365A-2  
Perfect score: 1080  
Sequence: 1 gggcgccgcccgtgcatg.....aaaaaaaaaaaaaaaa 1080

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pin.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1080	100.0	1080	6	AX211560	AX211560 Sequence
2	1080	100.0	1080	9	AF053537	AF053537 Homo sapi
3	1073	99.4	1141	9	BC030128	BC030128 Homo sapi
4	1046	96.9	1078	6	AR213555	AR213555 Sequence
5	1046	96.9	1078	9	AB012853	AB012853 Homo sapi
6	839.8	77.8	843	9	HSAJ76851	AJ006851 Homo sapi
7	838.4	77.6	840	6	AR213554	AR213554 Sequence
8	825.2	76.4	1154	6	AR220840	AR220840 Sequence
9	817.2	75.7	183317	9	HSING282	AF062748 Homo sapi
10	817.2	75.7	183317	9	AC107214	AC107214 Homo sapi
11	761.4	70.5	1001	10	AF078834	AF078834 Mus muscu
12	700.2	64.8	1262	10	BC050003	BC050003 Mus muscu
13	686.8	63.6	231311	2	AC132966	AC132966 Rattus no
14	686.8	63.6	245104	2	AC128348	AC128348 Rattus no
15	686.8	63.6	247796	2	AC094539	AC094539 Rattus no
16	686.8	63.6	255408	2	AC129658	AC129658 Rattus no
17	673.6	62.4	192753	2	AC107236	AC107236 Mus muscu
18	574.6	53.2	42029	2	AC114762	AC114762 Homo sapi
19	468	43.3	666	4	AB098919	AB098919 Bos tauru
20	432.8	40.1	1094	5	AY014017	AY014017 Xenopus l
21	332.2	30.8	346	6	BD052204	BD052204 Sequence
22	256.8	23.8	840	10	AF149820	AF149820 Mus muscu
23	253.6	23.5	1584	10	BC016573	BC016573 Mus muscu
24	253.6	23.5	2817	10	AF177757	AF177757 Mus muscu
25	239.2	22.1	423	9	HSING251	AF062747 Homo sapi
26	237.4	22.0	873	6	AR171884	AR171884 Sequence
27	237.4	22.0	911	9	AF044076	AF044076 Homo sapi
28	237.4	22.0	1533	9	AF149721	AF149721 Homo sapi
29	237.4	22.0	2444	9	AB024401	AB024401 Homo sapi
30	236	21.9	840	9	AF078835	AF078835 Homo sapi
31	236	21.9	845	9	HS310392	AJ310392 Homo sapi
32	236	21.9	1533	6	AX367042	AX367042 Sequence
33	236	21.9	2897	9	AF181850	AF181850 Homo sapi
34	207	19.2	201302	2	AC120896	AC120896 Rattus no
35	207	19.2	222955	2	AC135534	AC135534 Rattus no
36	199	18.4	184918	2	AC114608	AC114608 Mus muscu
37	198	18.4	235079	2	AC124475	AC124475 Mus muscu
38	198.8	18.4	1859	10	AF177755	AF177755 Mus muscu
39	197.8	18.3	1944	10	AF177756	AF177756 Mus muscu
40	192	17.8	222331	2	AC096794	AC096794 Rattus no
41	184.4	17.1	1980	9	AB024404	AB024404 Homo sapi
42	183.8	17.0	1074	9	AF149722	AF149722 Homo sapi
43	183.8	17.0	1143	6	AX367043	AX367043 Sequence
44	183.8	17.0	1902	6	AR079046	AR079046 Sequence
45	183.8	17.0	1902	6	AR087457	AR087457 Sequence

ALIGNMENTS

RESULT 1  
AX211560  
LOCUS AX211560 1080 bp DNA linear PAT 06-SEP-2001  
DEFINITION Sequence 7 from Patent WO0159114.  
ACCESSION AX211560  
VERSION AX211560.1 GI:15523814  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Harris,C.C. and Nagashima,M.  
TITLE Tumour suppressor Gene, p47ing3  
JOURNAL Patent: WO 0159114-A 7 16-AUG-2001;

Pred. No. is the number of results predicted by chance to have a

## THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

```
FEATURES
  source
    Location/Qualifiers
      1. .1080
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /notes="p33ING2"
BASE COUNT      375 a   196 c   271 g   238 t
ORIGIN
Query Match      100.0%; Score 1080; DB 6; Length 1080;
Best Local Similarity 100.0%; Pred. No. 1.2e-221;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCCGCGCGCGTGCATGTGCGGCTGCTGGATCGGAGCGCGCGCGCGCGGAT 60
DB 1 GCGGCCGCGCGCGTGCATGTGCGGCTGCTGGATCGGAGCGCGCGCGCGCGGAT 60
QY 61 CGGCAGGATGTTAGGCGCAGCAGCAGCAACTGTACTCGTCGGCTCGCTCTGACCCG 120
DB 61 CGGCAGGATGTTAGGCGCAGCAGCAGCAACTGTACTCGTCGGCTCGCTCTGACCCG 120
QY 121 GGAGCGAGCGCGCTGCTCACTGCTAGCTGTCAGGACTACCTTGTAGTGGTGGAGTCGCT 180
DB 121 GGAGCGAGCGCGCTGCTCACTGCTAGCTGTCAGGACTACCTTGTAGTGGTGGAGTCGCT 180
QY 181 GCCCAGCAGCATCCAGAGCACTGTCTGTCTGCGAGCTGCGACACAAATATCAAGA 240
DB 181 GCCCAGCAGCATCCAGAGCACTGTCTGTCTGCGAGCTGCGACACAAATATCAAGA 240
QY 241 AACGTTAAAGAAATGTATGATGCTCTACGAAATATAGAAAGAGATGATTTAAACCA 300
DB 241 AACGTTAAAGAAATGTATGATGCTCTACGAAATATAGAAAGAGATGATTTAAACCA 300
QY 301 GAAGAAACGCTCAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 360
DB 301 GAAGAAACGCTCAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 360
QY 361 TGAAGAAATACAGATTGTTACAAATGCTCGAATTTGGTGGAAATCGGCAAGACAAAT 420
DB 361 TGAAGAAATACAGATTGTTACAAATGCTCGAATTTGGTGGAAATCGGCAAGACAAAT 420
QY 421 GGAAGTTACACTCAGAGTGTTCAGAGTCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 480
DB 421 GGAAGTTACACTCAGAGTGTTCAGAGTCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 480
QY 481 AAAGATGGATTCCAGCCCAACAGAAAGATCTTCAAGAAAGACCCCGCAGCGGACCCAG 540
DB 481 AAAGATGGATTCCAGCCCAACAGAAAGATCTTCAAGAAAGACCCCGCAGCGGACCCAG 540
QY 541 TGAAGCGGTGATTTATGTACATGGCAATGGGATTTGAAGCTGTGATGATCAGCCACC 600
DB 541 TGAAGCGGTGATTTATGTACATGGCAATGGGATTTGAAGCTGTGATGATCAGCCACC 600
QY 601 TAAAGAAAGAAATCAAAGTCAGCAAGAAAGAAAGCGCTCCAAGCCCAAGCAGGAAAG 660
DB 601 TAAAGAAAGAAATCAAAGTCAGCAAGAAAGAAAGCGCTCCAAGCCCAAGCAGGAAAG 660
QY 661 GGAAGCTTCACTGTTGAGTTGCAATAGATCTCTAATGAACCTACACTGCTTTATGCAA 720
DB 661 GGAAGCTTCACTGTTGAGTTGCAATAGATCTCTAATGAACCTACACTGCTTTATGCAA 720
QY 721 CCAAGTGTCTTATGGGAGATGATAGATGTGCAATGAACAGTGTCCAATTTGAATGGTT 780
DB 721 CCAAGTGTCTTATGGGAGATGATAGATGTGCAATGAACAGTGTCCAATTTGAATGGTT 780
QY 781 TCACCTTTTCATGTTTTCACCTACCTATTAACCAAGGGAATGTTATTTGCCCAAGTG 840
DB 781 TCACCTTTTCATGTTTTCACCTACCTATTAACCAAGGGAATGTTATTTGCCCAAGTG 840
QY 841 CAGGGAGATAATGAGAAACAAATGGAACAAAGTACTGAAAGACAAAGGATAGAAG 900
DB 841 CAGGGAGATAATGAGAAACAAATGGAACAAAGTACTGAAAGACAAAGGATAGAAG 900
```

```
QY 901 ATCAGGAGTAGTAAGGCCATCCACATTTTAAAGGGTTATTTCTTTTATATAATTCGTT 960
DB 901 ATCAGGAGTAGTAAGGCCATCCACATTTTAAAGGGTTATTTCTTTTATATAATTCGTT 960
QY 961 TGCCTTTTCAGAAATGTTTTAGGGTAAATGCATAAGACTATGCAATAATTTTAAATCA 1020
DB 961 TGCCTTTTCAGAAATGTTTTAGGGTAAATGCATAAGACTATGCAATAATTTTAAATCA 1020
QY 1021 GTATTAAATGGTGTATTAAGTTGTTTACCTTTGAAAAAAGAAAAAAGAAAAA 1080
DB 1021 GTATTAAATGGTGTATTAAGTTGTTTACCTTTGAAAAAAGAAAAAAGAAAAA 1080

RESULT 2
AF053537      1080 bp      mRNA      linear      PRI 15-AUG-2001
LOCUS      Homo sapiens p33 (ING2) mRNA, complete cds.
DEFINITION      AF053537
ACCESSION      AF053537
VERSION      AF053537.1 GI:9992837
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1080)
AUTHORS      Nagashima,M., Shiseki,M., Miura,K., Hagiwara,K., Linke,S.P.,
Pedoux,R., Wang,X.W., Yokota,J., Riabowol,K. and Harris,C.C.
TITLE      DNA damage-inducible gene p33ING2 negatively regulates cell
JOURNAL      proliferation through acetylation of p53
MEDLINE      Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9671-9676 (2001)
PUBMED      21396501
11481424
REFERENCE      2 (bases 1 to 1080)
AUTHORS      Nagashima,M., Hagiwara,K., Minter,A.R. and Harris,C.C.
TITLE      Direct Submission
JOURNAL      Submitted (12-MAR-1998) Laboratory of Human Carcinogenesis,
National Cancer Institute, 37 Convent Drive Bldg.37 Rm.2C01,
Bethesda, MD 20892, USA
FEATURES
  source
    Location/Qualifiers
      1. .1080
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /chromosome="4"
      /map="4q35"
      1. .1080
      /gene="ING2"
      68. .910
      /gene="ING2"
      /codon_start=1
      /product="p33"
      /protein_id="AAG11395.1"
      /db_xref="GI:9992838"
      /translation="MLGQOQOOLYSALLTGERSLLTCTVQVDLECVESLPHDMQR
      NVSLRELNDKVKYQETLKEIDVYKKEDDLNKKRLQQLLQALRLINSQELGDEKIQ
      IYQMLELVENRAREMELHSCQFQDPASERASDKAKMDSSPERSRRPRQRSES
      RDLCHMANGIEDCDQPPKREKSKAKKREASPEVAIDNPEFYCLCN
      QVSYGEMIGCDNEQCPIEFHFPSCVSLTYKPKGWYCPKCRGDNKMTDKTEKTKD
      RRSR"
      106
      /gene="ING2"
      /note="polymorphism"
      /replace="c"
BASE COUNT      375 a   196 c   271 g   238 t
ORIGIN
Query Match      100.0%; Score 1080; DB 9; Length 1080;
Best Local Similarity 100.0%; Pred. No. 1.2e-221;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCCGCGCGCGTGCATGTGCGGCTGCTGGATCGGAGCGCGCGCGCGGAT 60
```



Db 69 GCGGCGCGCGGCTGCATGTGCGGCTGCTGGATCGGAGCGCGCGACGCGCGGAT 128  
Qy 61 CGGCAGGATGTTAGGCGAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTGACCGG 120  
Db 129 CGGCAGGATGTTAGGCGAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTGACCGG 188  
Qy 121 GGAGCGGAGCGGCTGCTCACTGCTACCTGCTGAGGACTTACCTTGTAGTGGTGGAGTCGCT 180  
Db 189 GGAGCGGAGCGGCTGCTCACTGCTGAGGACTTACCTTGTAGTGGTGGAGTCGCT 248  
Qy 181 GCCCCACGACATGACAGAGAAAGTGTGTGTGTCGAGAGCTGGACAAATAATCAAGA 240  
Db 249 GCCCCACGACATGACAGAGAAAGTGTGTGTGTCGAGAGCTGGACAAATAATCAAGA 308  
Qy 241 AACGTTAAAGGAAATGATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCA 300  
Db 309 AACGTTAAAGGAAATGATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCA 368  
Qy 301 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 360  
Db 369 GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGA 428  
Qy 361 TGAATAAATAACAGATGTTTACACAAATGCTCGAAATGTTGGTGGAAATCGGGCAAGACAAAT 420  
Db 429 TGAATAAATAACAGATGTTTACACAAATGCTCGAAATGTTGGTGGAAATCGGGCAAGACAAAT 488  
Qy 421 GGAGTTACACTACAGTGTGTTTCAAGATGCTGCTGAAAGTGAACGAGCCCTCAGATAAAGC 480  
Db 489 GGAGTTACACTACAGTGTGTTTCAAGATGCTGCTGAAAGTGAACGAGCCCTCAGATAAAGC 548  
Qy 481 AAGATGGATTCAGGCAACCCAGAAAGATCTTCAAGAAAGCCCGCAGCAGCGGACCAAG 540  
Db 549 AAGATGGATTCAGGCAACCCAGAAAGATCTTCAAGAAAGCCCGCAGCAGCGGACCAAG 608  
Qy 541 TGAAGCCGCTGATTTATGTACATGCGCAATGGGATTTGAAGATGTGTGATGATCAGCCACC 600  
Db 609 TGAAGCCGCTGATTTATGTACATGCGCAATGGGATTTGAAGATGTGTGATGATCAGCCACC 668  
Qy 601 TAAAGAAAGAAATCCAAAGTACAGCAAGAAAGAAACCGTCCAAAGCCCAAGCAGAAAG 660  
Db 669 TAAAGAAAGAAATCCAAAGTACAGCAAGAAAGAAACCGTCCAAAGCCCAAGCAGAAAG 728  
Qy 661 GGAAGCTTCACCTGTTGAGTTGCAATAGATCTTAATGAACCTACATCTGCTTATGCAA 720  
Db 729 GGAAGCTTCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATCTGCTTATGCAA 788  
Qy 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTTGAATGGTT 780  
Db 789 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTTGAATGGTT 848  
Qy 781 TCACCTTTTCATGTTTCACTTACCTTATAAACCAAGGGGAAATGGTATTTGCCCAAAGTG 840  
Db 849 TCACCTTTTCATGTTTCACTTACCTTATAAACCAAGGGGAAATGGTATTTGCCCAAAGTG 908  
Qy 841 CAGGGAGATTAATGAGAAACATGCAAAAGTCTGAAAGACAAAAGAGATAGAAG 900  
Db 909 CAGGGAGATTAATGAGAAACATGCAAAAGTCTGAAAGACAAAAGAGATAGAAG 968  
Qy 901 ATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTTATTTGTCTTTTATATAATTCGTT 960  
Db 969 ATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTTATTTGTCTTTTATATAATTCGTT 1028  
Qy 961 TGCTTTTCAGAAATGTTTAGGGTAAATGATGAAGCTATGCAATATTTTATCATTA 1020  
Db 1029 TGCTTTTCAGAAATGTTTAGGGTAAATGATGAAGCTATGCAATATTTTATCATTA 1088  
Qy 1021 GTATTAATGTTTAAAGTTGTTGCTACTTTTGAATAAATAAATAAATAAATAAATAAATAA 1073  
Db 1089 GTATTAATGTTTAAAGTTGTTGCTACTTTTGAATAAATAAATAAATAAATAAATAAATAA 1141

LOCUS AR213555 1078 bp DNA linear PAT 25-SEP-2002  
DEFINITION Sequence 7 from patent US 6403785.  
ACCESSION AR213555  
VERSION AR213555.1 GI:23310822  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1078)  
AUTHORS Nagata,M., Ozaki,K., Shimada,Y. and Horie,M.  
TITLE Isolated DNA molecule encoding human TSC403  
JOURNAL Patent: US 6403785-A 7 11-JUN-2002;  
FEATURES  
Location/Qualifiers  
1..1078  
source /organism="unknown"  
BASE COUNT 356 a 206 c 280 g 236 t  
ORIGIN  
Query Match 96.9%; Score 1046; DB 6; Length 1078;  
Best Local Similarity 99.5%; Pred. No. 2.4e-214;  
Matches 1049; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 GCGGCGCGCGCGGCTGCATGTGCGCTGCTCGATCGGAGCGCGCGGACGCGCGGAT 60  
Db 25 GCGGCGCGCGCGGCTGCATGTGCGCTGCTCGATCGGAGCGCGCGGACGCGCGGAT 84  
Qy 61 CGGCAGGATGTTAGGCGAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTGACCGG 120  
Db 85 CGGCAGGATGTTAGGCGAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTGACCGG 144  
Qy 121 GGAGCGGAGCGGCTGCTCACCTGCTGAGGACTTACCTTGTAGTGGTGGAGTCGCT 180  
Db 145 GGAGCGGAGCGGCTGCTCACCTGCTGAGGACTTACCTTGTAGTGGTGGAGTCGCT 204  
Qy 181 GCCCCACGACATGACAGAGAAAGTGTGTGTGTCGAGAGCTGGACAAATAATCAAGA 240  
Db 205 GCCCCACGACATGACAGAGAAAGTGTGTGTGTCGAGAGCTGGACAAATAATCAAGA 264  
Qy 241 AACGTTAAAGGAAATGATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCA 300  
Db 265 AACGTTAAAGGAAATGATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCA 324  
Qy 301 GAAGAAACGCTTACAGCAGCTTCTCCAGAGACACTTAATTAATAGTCAAGAAATGGGAGA 360  
Db 325 GAAGAAACGCTTACAGCAGCTTCTCCAGAGACACTTAATTAATAGTCAAGAAATGGGAGA 384  
Qy 361 TGAATAAATAACAGATGTTTACACAAATGCTCGAAATGTTGGTGGAAATCGGGCAAGACAAAT 420  
Db 385 TGAATAAATAACAGATGTTTACACAAATGCTCGAAATGTTGGTGGAAATCGGGCAAGACAAAT 444  
Qy 421 GGAGTTACACTACAGTGTGTTTCCAAAGATCTGCTGAAAGTGAAAGAGCCCTCAGATAAAGC 480  
Db 445 GGAGTTACACTACAGTGTGTTTCCAAAGATCTGCTGAAAGTGAAAGAGCCCTCAGATAAAGC 504  
Qy 481 AAGATGAGTTCCAGCCCAACCAAGAGATCTTCAAGAGAGACCCCGCAGCAGCGGACCAAG 540  
Db 505 AAGATGAGTTCCAGCCCAACCAAGAGATCTTCAAGAGAGACCCCGCAGCAGCGGACCAAG 564  
Qy 541 TGAAGCGCTGATTTATGTTCATGCGCAAAATGGGATTTGAAGACTGTGATGATCAGCCACC 600  
Db 565 TGAAGCGCTGATTTATGTTCATGCGCAAAATGGGATTTGAAGACTGTGATGATCAGCCACC 624  
Qy 601 TAAAGAAAGAAATCCAAAGTACGAAAGAAAGAAAGAGAGCTCCAAAGGCCAAGCAGAAAG 660  
Db 625 TAAAGAAAGAAATCCAAAGTACGAAAGAAAGAAAGAGAGCTCCAAAGGCCAAGCAGAAAG 684  
Qy 661 GGAAGCTTCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATCTGCTTATGCAA 720  
Db 685 GGAAGCTTCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATCTGCTTATGCAA 744  
Qy 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTTGAATGGTT 780  
Db 745 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTTGAATGGTT 804



```
QY 1021 GTATTAAATGCTGTTATTAAGTTGTTGTTACTTTG 1054
Db 1045 GTATTAAATGCTGTTATTAAGTTGTTGTTACTTTG 1078

RESULT 6
HSAJ6851 843 bp mRNA linear PRI 06-JAN-2001
LOCUS Homo sapiens mRNA for p32 protein.
ACCESSION AJ006851
VERSION AJ006851.1 GI:12053587
KEYWORDS ing2 gene; p32 protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Cal.S., Freije, J.M. and Lopez-Otin, C.
ING2, a new possible gene supressor tumor
Unpublished
2 (bases 1 to 843)
Lopez-Otin, C.
Direct Submission
Submitted (08-JUN-1998) Lopez-Otin C., Bioquimica y Biologia
Molecular, Universidad de Oviedo. Facultad de Medicina., C/ Julian
Claveria s/n., 33006, SPAIN
FEATURES
Location/Qualifiers
1..843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="breast tumor"
1..843
/gene="ing2"
1..843
/gene="ing2"
/codon_start=1
/product="p32 protein"
/protein_id="CAC20567.1"
/db_xref="GI:12053588"
/translation="MLGQQQQLYSSAALLTGERTRLTLTYVQDYLVCESLPHDMQR
IVSYLRELDNKKYQETLKEIDVYKVEDDLNOKRLQQLRALINSOELGDEKIO
NVTOMLELVENAROMELHSQCFQPAESERADKAMDSDOPERSRRPRQTSSES
RDLCHMANGIEDCDQPPKPKKSKAKKKRSKAKQERASPVFAIDNBEPTCLCN
QVSYGEMIGCDNEQCEPIEFHFSCVSLTYPKGKWCPKCRGDNBKTMDKSTKTKKD
RRSR"
BASE COUNT 298 a 163 c 212 g 170 t
ORIGIN
Query Match 77.8%; Score 839.8; DB 9; Length 843;
Best Local Similarity 99.8%; Pred. No. 4.8e-170;
Matches 841; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 68 ATGTAGGCGCAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTCTGACCGGGAGCGG 127
Db 1 ATGTAGGCGCAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTCTGACCGGGAGCGG 60
QY 128 AGCCGGCTGCTCACCTGTCTACGTGCAGGACTACCTTTGAGTGGGTGGAGTGCCTGCCCCAC 187
Db 61 ACGGGCTGCTCACCTGTCTACGTGCAGGACTACCTTTGAGTGGGTGGAGTGCCTGCCCCAC 120
QY 188 GACATGCAGAGGAACCTGTCTGTGTCGCGAGAGCTGGACAAACAATATCAAGAAACGTTA 247
Db 121 GACATGCAGAGGAACCTGTCTGTGTCGCGAGAGCTGGACAAACAATATCAAGAAACGTTA 180
QY 248 AAGGAAATTTGATGATCTCTACGAAAAATATAAGAAAGAGATGATTTTAAACCAGAGAAA 307
Db 181 AAGGAAATTTGATGATCTCTACGAAAAATATAAGAAAGAGATGATTTTAAACCAGAGAAA 240
QY 308 CGTCTACAGCAGCTTCTCCAGAGAGCAGCACTAATTAATAGTCAAGAAATGGGAGATGAAAA 367
Db 241 CGTCTACAGCAGCTTCTCCAGAGAGCAGCACTAATTAATAGTCAAGAAATGGGAGATGAAAA 300
```

```
QY 368 ATACAGATTGTTACAAAATGCTCGAATTGGTGGAAAAATCGGGCAAGACAAAATGAGTTA 427
Db 301 ATACAGATTGTTACAAAATGCTCGAATTGGTGGAAAAATCGGGCAAGACAAAATGAGTTA 360
QY 428 CACTCACAGTGTTCCTCAAGATCTCTGCTGAAAGTGAACGAGCTCTCAGATAAAGCAAGATG 487
Db 361 CACTCACAGTGTTCCTCAAGATCTCTGCTGAAAGTGAACGAGCTCTCAGATAAAGCAAGATG 420
QY 488 GATTCACGCCACCAAGAAAGATCTTCAAGNAGACCCCGCAGCAGCGGACCAAGTGAAGC 547
Db 421 GATTCACGCCACCAAGAAAGATCTTCAAGNAGACCCCGCAGCAGCGGACCAAGTGAAGC 480
QY 548 CGTGAATTTATGTGCATGSCAAATGGGATTGAAGACTGTGTGATGATCAGCCACCCTAAAGAA 607
Db 481 CGTGAATTTATGTGCATGSCAAATGGGATTGAAGACTGTGTGATGATCAGCCACCCTAAAGAA 540
QY 608 AAGAAATCCAAGTCAGCAAGAAAAGAAACGCTCCAAGGCCAAGCAGCAAGAAAGGAGCT 667
Db 541 AAGAAATCCAAGTCAGCAAGAAAAGAAACGCTCCAAGGCCAAGCAGCAAGAAAGGAGCT 600
QY 668 TCACCTGTTGAGTTTGCAATAGATCCTTAATGAACCTACATACCTTATGCAACCAAGTG 727
Db 601 TCACCTGTTGAGTTTGCAATAGATCCTTAATGAACCTACATACCTTATGCAACCAAGTG 660
QY 728 TCTTATGGGAGATGATAGGATGTGCAATGAACAGTGTCCAAATTTGAATGGTTTCACTTT 787
Db 661 TCTTATGGGAGATGATAGGATGTGCAATGAACAGTGTCCAAATTTGAATGGTTTCACTTT 720
QY 788 TCATGTGTTTCACTTTACCTATAAACCAGGGGAAATGGTATTTGCCCAAAGTCAGGGGA 847
Db 721 TCATGTGTTTCACTTTACCTATAAACCAGGGGAAATGGTATTTGCCCAAAGTCAGGGGA 780
QY 848 GATAATGAGAAAACAATGCAAAAAGTACTGAAAAGACAAAAGATAGAAGATCGAGG 907
Db 781 GATAATGAGAAAACAATGCAAAAAGTACTGAAAAGACAAAAGATAGAAGATCGAGG 840
QY 908 TAG 910
Db 841 TAG 843

RESULT 7
LOCUS AR213554
DEFINITION Sequence 6 from patent US 6403785.
ACCESSION AR213554
VERSION AR213554.1 GI:23310821
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 840)
AUTHORS Nagata,M., Ozaki,K., Shimada,Y. and Horie,M.
TITLE Isolated DNA molecule encoding human TSC403
JOURNAL Patent: US 6403785-A, 6 11-JUN-2002;
FEATURES Location/Qualifiers
1..840
/organism="unknown"
BASE COUNT 297 a 164 c 211 g 168 t
ORIGIN
Query Match 77.6%; Score 838.4; DB 6; Length 840;
Best Local Similarity 99.9%; Pred. No. 9.6e-170;
Matches 839; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 68 ATGTAGGCGCAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTCTGACCGGGAGCGG 127
Db 1 ATGTAGGCGCAGCAGCAGCAACTGTACTCGTGGCTGCGCTCTCTGACCGGGAGCGG 60
QY 128 AGCCGGCTGCTCACCTGTCTACGTGCAGGACTACCTTTGAGTGGGTGGAGTGCCTGCCCCAC 187
Db 61 AGCCGGCTGCTCACCTGTCTACGTGCAGGACTACCTTTGAGTGGGTGGAGTGCCTGCCCCAC 120
```

```
QY 188 GACATGAGAGGAAACGCTGCTGCTGAGAGCTGAGACAAACAAATATCAAGAAACGTTA 247
Db 121 GACATGAGAGGAAACGCTGCTGCTGAGAGCTGAGACAAACAAATATCAAGAAACGTTA 180
QY 248 AAGGAAATTCATGATGCTACGAAATATTAAGAAAGAGATGATTTAAACCAAGAGAA 307
Db 181 AAGGAAATTCATGATGCTACGAAATATTAAGAAAGAGATGATTTAAACCAAGAGAA 240
QY 308 CGTCTACAGAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGATGAA 367
Db 241 CGTCTACAGAGCTTCTCCAGAGAGCACTAAATTAATAGTCAAGAAATGGGAGATGAA 300
QY 368 ATACAGATTTTACAAATGCTCGAATTCGATTCGAAATCGGCAAGACAAATGGAGTTA 427
Db 301 ATACAGATTTTACAAATGCTCGAATTCGATTCGAAATCGGCAAGACAAATGGAGTTA 360
QY 428 CACTCAGAGTTCCTCAAGATCTCTGTAAGTGAAGAGCAAGCTCAGATAAGCAAGATG 487
Db 361 CACTCAGAGTTCCTCAAGATCTCTGTAAGTGAAGAGCAAGCTCAGATAAGCAAGATG 420
QY 488 GATTCCAGCAACCCAGAAAGATCTTCAAGAGACCCCGCAGGAGCGGACCACTGAAAGC 547
Db 421 GATTCCAGCAACCCAGAAAGATCTTCAAGAGACCCCGCAGGAGCGGACCACTGAAAGC 480
QY 548 CGTGATTTATGTCATGCGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAA 607
Db 481 CGTGATTTATGTCATGCGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAA 540
QY 608 AAGAAATCCAGTCAGCAAAAGAAAGAAACGCTCCAGGCCCAAGCAAGGAAAGGAGCT 667
Db 541 AAGAAATCCAGTCAGCAAAAGAAAGAAACGCTCCAGGCCCAAGCAAGGAAAGGAGCT 600
QY 668 TCACCTGTTGAGTTTGAATAGATCCTTAATGAACCTACATCTGCTTATGCAACCAAGTG 727
Db 601 TCACCTGTTGAGTTTGAATAGATCCTTAATGAACCTACATCTGCTTATGCAACCAAGTG 660
QY 728 TCTTATGGGAGATGATAGGATGTGCAATGAACAGTGTCCAATGAAATGGTTTCACTTT 787
Db 661 TCTTATGGGAGATGATAGGATGTGCAATGAACAGTGTCCAATGAAATGGTTTCACTTT 720
QY 788 TCATGTTGTTCACTTACCTATAAACCAAGGGGAAATGGTATTGCCCAAGTCAGGGGA 847
Db 721 TCATGTTGTTCACTTACCTATAAACCAAGGGGAAATGGTATTGCCCAAGTCAGGGGA 780
QY 848 GATAATGAGAAACAAATGGCAAAAGTACTGAAAGACAAAGAGATGAAGATCGAGG 907
Db 781 GATAATGAGAAACAAATGGCAAAAGTACTGAAAGACAAAGAGATGAAGATCGAGG 840

RESULT 8
AR220840
LOCUS AR220840 1154 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 81 from patent US 6426186.
ACCESSION AR220840
VERSION AR220840.1 GI:23327717
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1154)
AUTHORS Jones, K.A., Volkmut, W. and Walker, M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 81 30-JUL-2002;
FEATURES Location/Qualifiers
1..1154
/organism="unknown"
BASE COUNT 367 a 216 c 285 g 265 t 21 others
ORIGIN
```

Query Match 76.4%; Score 825.2; DB 6; Length 1154;  
Best Local Similarity 97.9%; Pred. No. 7e-167;  
Matches 836; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
QY 211 GCTGCGAGAGCTGGACAAACAAATATCAAGAAACGTTTAAAGGAAATTTGATGATCTACGA 270
Db 301 GCTGCGAGAGCTGGACAAACAAATATCAAGAAACGTTTAAAGGAAATTTGATGATCTACGA 360
QY 271 AAAATATTAAGAAAGAGATGATTTAAACCAAGAAAGACGTCTACAGCAGCTTCTCCAGAG 330
Db 361 AAAATATTAAGAAAGAGATGATTTAAACCAAGAAAGACGTCTACAGCAGCTTCTCCAGAG 420
QY 331 AGCAGCTAAATTAATAGTCAAGAAATGGGAGATGAAAGAAATACAGATTTGTTACAAATGCT 390
Db 421 AGCAGCTAAATTAATAGTCAAGAAATGGGAGATGAAAGAAATACAGATTTGTTACAAATGCT 480
QY 391 CGAATTTGTTGAAAAATCGGGCAAGACAAATGGAGTTTACACTCACAGTGTTTCCAGATCC 450
Db 481 CGAATTTGTTGAAAAATCGGGCAAGACAAATGGAGTTTACACTCACAGTGTTTCCAGATCC 540
QY 451 TCGTGAAGTGAACAGAGCTCAGATAAGCAAGATGGATTTCCAGCAACCAAGAAAGATC 510
Db 541 TCGTGAAGTGAACAGAGCTCAGATAAGCAAGATGGATTTCCAGCAACCAAGAAAGATC 600
QY 511 TTCAAGAAAGACCCCGCAGGAGCGGACCACTGAAAGCCGCTGATTTATGTCACATGGCAA 570
Db 601 TTCAAGAAAGACCCCGCAGGAGCGGACCACTGAAAGCCGCTGATTTATGTCACATGGCAA 660
QY 571 TGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAAGTCAGCAAGAA 630
Db 661 TGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAAGTCAGCAAGAA 720
QY 631 AAGAAACCTCTCAAGGCCCAAGCAAGGAAAGGAGCTTCACTGTTGAGTTTGCATATAGA 690
Db 721 AAGAAACCTCTCAAGGCCCAAGCAAGGAAAGGAGCTTCACTGTTGAGTTTGCATATAGA 780
QY 691 TCCTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATGAGTG 750
Db 781 TCCTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATGAGTG 840
QY 751 TGCAATGAACAGTGTCCAAATTTGAATGGTTTCACTTTTCATGTTTTCATCTTACCTATATA 810
Db 841 TGCAATGAACAGTGTCCAAATTTGAATGGTTTCACTTTTCATGTTTTCATCTTACCTATATA 900
QY 811 ACCAAAGGGGAAATGGTATTGCCCAAGTGCAGGGGAGATATAGAAAAACAATGGACAA 870
Db 901 ACCAAAGGGGAAATGGTATTGCCCAAGTGCAGGGGAGATATAGAAAAACAATGGACAA 960
QY 871 AAGTACTGAAAGACAAAGAAAGATAGAGATCGAGTGTAGTAAGGCCATCCACATTTTA 930
Db 961 AAGTACTGAAAGACAAAGAAAGATAGAGATCGAGTGTAGTAAGGCCATCCACATTTTA 1020
QY 931 AAGGGTTATTGCTCTTTTATATAATTCGTTTTCCTTTTCAGAAAAATGTTTAGGGTAAATGC 990
Db 1021 AAGGGTTATTGCTCTTTTATATAATTCGTTTTCCTTTTCAGAAAAATGTTTAGGGTAAATGC 1080
QY 991 ATAAGACTATGCAATAATTTTAAATCATTAGTATTAAATCGTGTATTAAAGTTGTTGTAC 1050
Db 1081 ATAAGACTATGCAATAATTTTAAATCATTAGTATTAAATCGTGTATTAAAGTTGTTGTAC 1140
QY 1051 TTTGAAAAAATAA 1064
Db 1141 TTTGAAAAAATAA 1154

RESULT 9
HSING2S2
LOCUS HSING2S2 974 bp DNA linear PRI 08-SEP-2000
DEFINITION Homo sapiens p33 (ING2) gene, exon 2, complete sequence; and complete cds.
ACCESSION AF062748
VERSION AF062748.1 GI:9992841
KEYWORDS
SEGMENT 2 of 2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 974)

REFERENCE Nagashima, M., Hagiwara, K., Minter, A.R. and Harris, C.C. 1998. Direct Submission. Submitted (01-MAY-1998) Laboratory of Human Carcinogenesis, National Cancer Institute, 37 Convent Dr. Bldg. 37, Rm. 2C26, Bethesda, MD 20892, USA

FEATURES

source 1. .974

gene /organism="Homo sapiens" /mol\_type="genomic DNA" /db\_xref="taxon:9606" order(AF062747.1:1. .423,1. .938)

mRNA /genes="ING2" /join(AF062747.1:1. .239,124. .938)

CDS /product="p33 ING2" /join(AF062747.1:68. .239,124. .794)

/gene="ING2" /codon\_start=1 /product="p33" /protein\_id="AAG11396.1" /db\_xref="GI:992842" /translation="MLGQQOOLYSALLTGRSRLTTCYVDYLECVESLPHDMOR NVSVRELNKKYQETLKEIDDDYKYEKKEDDLNKKRLQQLQRLINQSLGDEKIQ IVQMLVELVENRQARQELHSCQFQPAESRSDKAKMDSQSPERSRPRRORTSES RDLCHWANGIEDCDDQFPKPEKKSAAKKRKAQERASVPFEAIDNPEPTCYLCN QVSYGEMIGCDNEQCPLEWFFHFCVSLYKPKGWYCPKRCGDNKMTDKSTKTKKD RRSR"

intron <1. 123 /gene="ING2" /number=1

exon 124. .938 /gene="ING2" /number=2

polya\_signal 918. 923 /gene="ING2" /evidence=experimental

BASE COUNT 348 a 156 c 205 g 265 t

ORIGIN

Query Match 75.7%; Score 817.2; DB 9; Length 974; Best Local Similarity 99.6%; Pred. No. 3.6e-165; Matches 819; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 233 TATCAAGAAACGTTAAAGGAAATGTATGATGTTCTACGAAAAATATAAGAAAGAGATGAT 292

DB 117 TTTTAAAGCGTTAAAGGAAATGTATGATGTTCTACGAAAAATATAAGAAAGAGATGAT 176

QY 293 TTAACCAAGAAACGTTACAGCAGCTTCTCCAGAGCAGCTAATTAATAGTCAAGAA 352

DB 177 TTAACCAAGAAACGTTACAGCAGCTTCTCCAGAGCAGCTAATTAATAGTCAAGAA 236

QY 353 TTGGGAGATCAAAAATACAGATGTTTACACAAATGCTCGAATTCGTGGAATCGGCA 412

DB 237 TTGGGAGATCAAAAATACAGATGTTTACACAAATGCTCGAATTCGTGGAATCGGCA 296

QY 413 AGACAAATGAGATTACACTCACAGTGTCTTCCAGATCCTGCTGAAAGTGAACGAGCTCA 472

DB 297 AGACAAATGAGATTACACTCACAGTGTCTTCCAGATCCTGCTGAAAGTGAACGAGCTCA 356

QY 473 GATAAGCAAGATGAGATTCCAGCCAAACGAAAGATCTTCAAGAGAGACCCCGAGCGAG 532

DB 357 GATAAGCAAGATGAGATTCCAGCCAAACGAAAGATCTTCAAGAGAGACCCCGAGCGAG 416

QY 533 CGGACCAAGTCAAGCCGCTGATTTATGTCATCGCAATCGGAATTCAGACTGTGATGAT 592

DB 417 CGGACCAAGTCAAGCCGCTGATTTATGTCATCGCAATCGGAATTCAGACTGTGATGAT 476

QY 593 CAGCCACCTAAAGAAAGAAATCCAAAGTCAGCAAGAAAGAAAGAAAGCTTCAAGGCCAAG 652

DB 477 CAGCCACCTAAAGAAAGAAATCCAAAGTCAGCAAGAAAGAAAGAAAGCTTCAAGGCCAAG 536

QY 653 CAGGAAAGGGAAGCTTCACTGTTGAGTTTGAATAGATCCCTAATGAACCTCATACTGC 712

DB 537 CAGGAAAGGGAAGCTTCACTGTTGAGTTTGAATAGATCCCTAATGAACCTCATACTGC 596

QY 713 TTATGCAACCAAGTGTCTTTATGGGAGATGATAGGATGTGACAATGAACAGTCCAAAT 772

DB 597 TTATGCAACCAAGTGTCTTTATGGGAGATGATAGGATGTGACAATGAACAGTCCAAAT 656

QY 773 GAATGGTTTCACTTTTTCATGTGTTTTCACCTTACTATAAACCAAGGGGAAATGTTATGC 832

DB 657 GAATGGTTTCACTTTTTCATGTGTTTTCACCTTACTATAAACCAAGGGGAAATGTTATGC 716

QY 833 CCAAAGTGCAGGGAGATAATCAGAAAAACAATGGACAAAGTACTCAAAAGACAAAAAAG 892

DB 717 CCAAAGTGCAGGGAGATAATCAGAAAAACAATGGACAAAGTACTCAAAAGACAAAAAAG 776

QY 893 GATAGAAGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTTATTTGCTTTTATAT 952

DB 777 GATAGAAGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTTATTTGCTTTTATAT 836

QY 953 AATTGCTTTGCTTTCAGAAAAATGTTTATAGGTAATGCAATAGACTATGCAATAATTTT 1012

DB 837 AATTGCTTTGCTTTCAGAAAAATGTTTATAGGTAATGCAATAGACTATGCAATAATTTT 896

QY 1013 AATCATTAGTATTAAATGGTGTATTAAGCTTGTGTACTTTG 1054

DB 897 AATCATTAGTATTAAATGGTGTATTAAGCTTGTGTACTTTG 938

RESULT 10

AC107214 183317 bp DNA linear PRI 15-OCT-2002

LOCUS Homo sapiens BAC clone RP11-367N14 from 4, complete sequence.

DEFINITION AC107214

ACCESSION AC107214.5 GI:23238103

VERSION

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS Nguyen, C., Haglund, K. and Spalding, L.

TITLE The sequence of Homo sapiens BAC clone RP11-367N14

JOURNAL Unpublished (2001)

MEDLINE 3 (bases 1 to 183317)

REFERENCE Waterston, R.H.

AUTHORS Direct Submission

JOURNAL Submitted (16-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 183317)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (10-APR-2002) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 183317)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (20-SEP-2002) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 6 (bases 1 to 183317)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (03-OCT-2002) Department of Genetics, Washington

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 7 (bases 1 to 183317)  
 Waterston,R.  
 Direct Submission  
 Submitted (15-OCT-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Sep 20, 2002 this sequence version replaced gi:20128734.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
 ----- Summary Statistics  
 -----  
 Center Project name: H\_NH0367N14  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
 VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC112698.

There is an unresolved tandem repeat from base 87660 to 89441.

Data from AC112698 was used to finish this clone.

#### FEATURES

source	Location/Qualifiers
	1. 183317
	/organism="Homo sapiens"
	/mol type="genomic DNA"
	/db xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-367N14"
	/clone lib="RPCI-11"
	66. 840
repeat_region	/rpt family="L1"
	856. 1214
repeat_region	/rpt family="MER1_type"
	2782. 3132
repeat_region	/rpt family="L1"
	3376. 4311
repeat_region	/rpt family="MER1_type"
	4312. 4642
repeat_region	/rpt family="ERV1"
	4643. 4762
repeat_region	/rpt family="Alu"
	4763. 5003
repeat_region	/rpt family="ERV1"
	5009. 5298
repeat_region	/rpt family="MER1_type"
	5299. 5605
repeat_region	/rpt family="Alu"
	5606. 5689
repeat_region	/rpt family="MER1_type"
	6690. 6920
repeat_region	/rpt family="Alu"
	6921. 7263
repeat_region	/rpt family="MER1_type"
	7405. 7452
repeat_region	/rpt family="L2"
	7793. 7864
repeat_region	/rpt family="MER1_type"
	8141. 8233
repeat_region	/rpt family="L2"
	8235. 8583
repeat_region	/rpt family="L1"
	8584. 8853
repeat_region	/rpt family="Alu"
	8854. 8933
repeat_region	/rpt family="L1"
	8953. 9114
repeat_region	/rpt family="Alu"
	9120. 9417
repeat_region	/rpt family="Alu"
	9422. 9836
repeat_region	/rpt family="Alu"
	9946. 10061
repeat_region	/rpt family="Alu"
	10696. 10958
repeat_region	/rpt family="Alu"
	10964. 11110
repeat_region	/rpt family="MER1_type"
	11580. 11886
repeat_region	/rpt family="Alu"
	12156. 12476
repeat_region	/rpt family="Alu"
	12488. 13030
repeat_region	/rpt family="ERV1"
	13077. 13192
repeat_region	/rpt family="(TTCC)n"
	13225. 13538
repeat_region	/rpt family="Alu"
	13634. 13749
repeat_region	/rpt family="MIR"
	14111. 14206
repeat_region	/rpt family="L1"
	14565. 14864
repeat_region	/rpt family="Alu"
	14982. 15277
repeat_region	/rpt family="Alu"
	15357. 15652
repeat_region	/rpt family="Alu"
	15709. 16001
repeat_region	/rpt family="Alu"
	16730. 16751
repeat_region	/rpt family="Alu"
	18157. 18271
repeat_region	/rpt family="Alu"
	19348. 19529
repeat_region	/rpt family="Alu"
	19874. 20012
repeat_region	/rpt family="L1"
	20068. 20103
repeat_region	/rpt family="AT-rich"
	20263. 20564
repeat_region	/rpt family="Alu"
	20769. 20995
repeat_region	/rpt family="Alu"
	21095. 21383

repeat\_region /rpt\_family="Alu"  
21605..21740  
repeat\_region /rpt\_family="Alu"  
21746..22060  
repeat\_region /rpt\_family="Alu"  
22072..22381  
repeat\_region /rpt\_family="Alu"  
22432..22526  
repeat\_region /rpt\_family="MIR"  
23055..23181  
repeat\_region /rpt\_family="Alu"  
23182..23212  
/rpt\_family="(TAA)n"

Query Match 75.7%; Score 817.2; DB 9; Length 183317;  
Best Local Similarity 99.8%; Pred. No. 1.1e-164;  
Matches 819; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 233 TATCAAGAAACGTTAAAGGAAATTTGATGTTCTACGAAATATATAGAAAGAGATGAT 292  
DB 45203 TTTTAGAACGTTAAAGGAAATTTGATGTTCTACGAAATATATAGAAAGAGATGAT 45262  
QY 293 TTAACACAGAGAAACGTTCTACAGCAGCTTCTCCAGAGACACTAATTAATAGTCAAGAA 352  
DB 45263 TTAACACAGAGAAACGTTCTACAGCAGCTTCTCCAGAGACACTAATTAATAGTCAAGAA 45322  
QY 353 TTGGGAGATGAAATAATACAGATTTTACACAATGCTCGAATTTGTTGNAATCGGCA 412  
DB 45323 TTGGGAGATGAAATAATACAGATTTTACACAATGCTCGAATTTGTTGNAATCGGCA 45382  
QY 413 AGACAAATGAGTTACACTCACAGTGTCTTCCAAAGTCTTCTGAAAGTGAACGAGCTCA 472  
DB 45383 AGACAAATGAGTTACACTCACAGTGTCTTCCAAAGTCTTCTGAAAGTGAACGAGCTCA 45442  
QY 473 GATAAAGCAAGATGAGTTCACGCCAACCAAGAAAGATCTTCAAGAGACACCCCGAGCAG 532  
DB 45443 GATAAAGCAAGATGAGTTCACGCCAACCAAGAAAGATCTTCAAGAGACACCCCGAGCAG 45502  
QY 533 CGACCAAGTGAAGCCGTTGATTTATGTCACATCGCAATGGGATTTGAAGCTGTGATGAT 592  
DB 45503 CGACCAAGTGAAGCCGTTGATTTATGTCACATCGCAATGGGATTTGAAGCTGTGATGAT 45562  
QY 593 CAGCCACCTAAAGAAAGAAATCCAGTCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAG 652  
DB 45563 CAGCCACCTAAAGAAAGAAATCCAGTCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAG 45622  
QY 653 CAGAAAGGAGACCTTCACTGTTGATTTGCAATAGATCCTAATGAACCTACATCTGC 712  
DB 45623 CAGAAAGGAGACCTTCACTGTTGATTTGCAATAGATCCTAATGAACCTACATCTGC 45682  
QY 713 TTATGCACCAAGTGTCTTATGGGAGATGATAGGATGACAAATGAACAGTGTCCAATT 772  
DB 45683 TTATGCACCAAGTGTCTTATGGGAGATGATAGGATGACAAATGAACAGTGTCCAATT 45742  
QY 773 GAATGTTTTCATTTTCACTGTTTCACTTTACCTATAAACCAAGGGAATGTTATTC 832  
DB 45743 GAATGTTTTCATTTTCACTGTTTCACTTTACCTATAAACCAAGGGAATGTTATTC 45802  
QY 833 CCAAGTGCAGGGGAGATTAATGAGAAACCAATGGACAAAGTACTGAAAGACAAAGAAAG 892  
DB 45803 CCAAGTGCAGGGGAGATTAATGAGAAACCAATGGACAAAGTACTGAAAGACAAAGAAAG 45862  
QY 893 GATAGAAGATCGAGGTAGTAAGGCCATCCACATTTTAAAGGGTATTTGCTCTTTATAT 952  
DB 45863 GATAGAAGATCGAGGTAGTAAGGCCATCCACATTTTAAAGGGTATTTGCTCTTTATAT 45922  
QY 953 AATTCGTTTCTCTTTCAGAAATGTTTTAGGGTAAATGCAATAGACTATGCAATAATTTTT 1012  
DB 45923 AATTCGTTTCTCTTTCAGAAATGTTTTAGGGTAAATGCAATAGACTATGCAATAATTTTT 45982  
QY 1013 AATCATTAGTATTAATGGTGTATTAAGATTTGTTGACTTTG 1054  
DB 45983 AATCATTAGTATTAATGGTGTATTAAGATTTGTTGACTTTG 46024

RESULT 11  
AF078834  
LOCUS 1001 bp mRNA linear ROD 09-SEP-2000  
DEFINITION Mus musculus p33ING2 (Ing2) mRNA, complete cds.  
ACCESSION AF078834  
VERSION AF078834.1 GI:10039542  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1001)  
AUTHORS Nagashima M., Hagiwara K., Hancock A.R. and Harris C.C.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUN-1998) Laboratory of Human Carcinogenesis,  
National Cancer Institute, National Institutes of Health, 37  
Convent Dr., Bldg. 37, Rm. 2C01, Bethesda, MD 20892, USA  
FEATURES  
Location/Qualifiers  
1..1001  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..1001  
/gene="Ing2"  
8..853  
/gene="Ing2"  
/codon\_start=1  
/product="p33ING2"  
/protein\_id="AAG12173.1"  
/db\_xref="GI:10039543"  
translation="MLGQOQOOLYSSALLTGBRSLLTCVVDVLECVESLPHDMQ  
RNVSLRELDNRYQETLEIDVDYKEYEKEDDSNOKRLQHLQALINSLQELDKSI  
QNVTLMLGVENRSQFMELHSQCFQDPAESRSKSDMSQPSRRSRPPQRTSE  
SRDLCHMTNGIDDDQPPKPKRSKAKKKKSKAKQEREASVFPADPNPEPTCLC  
NQVSYGEMIGCDNEQCPIEWFHFSCVSLTYKHKKWYCPKCRGDNKTKMDKSTKTK  
ERRAR"  
BASE COUNT 324 a 205 c 264 g 208 t  
Query Match 70.5%; Score 761.4; DB 10; Length 1001;  
Best Local Similarity 88.1%; Pred. No. 3.5e-153;  
Matches 887; Conservative 0; Mismatches 111; Indels 9; Gaps 5;  
QY 61 CGCAGGATGTTAGG---GCAGCAGCAGCAGCAACTGTACTCTGCGGCTCGGCTCTGAC 117  
DB 1 CGCAGGATGTTAGGCGCAGCAGCAGCAGCAGCTGTACTCTGCGGCGCGCTCTGAC 60  
QY 118 CGGGAGCGGAGCGGCTCTCACCTGTACGTGACGAGGACTACCTTGTAGTGCCTGAGTC 177  
DB 61 CGGAGAGCGGAGCGGCTCTCACCTGTACGTGACGAGGACTACCTGAGTGTGTGAGTC 120  
QY 178 GCTGCCCCACGACATGCGAGGAAACGTTCTGTGTCTGCGAGAGCTGAGCAACAAATATCA 237  
DB 121 GCTGCCCCACGACATGCGAGGAAACGTTCTGTGTCTGCGGAGCTGAGCAACAAATATCA 180  
QY 238 AGAAACGTTAAAGGAAATTTGATGATGTTCTACGAAATAATATAGAAAGAGATGATTAA 297  
DB 181 AGAAACGTTAAAGGAAATTTGATGATGTTCTATGAAAAATATGAGAAAGAGATGATTCCAA 240  
QY 298 CCAGAGAAACGCTTACAGCAGCTTCTCCAGAGAGCCTTAATTAATAGTCAAGATTTGG 357  
DB 241 CCAGAAACGCTTACAGCAGCTTCTCCAGAGAGCCTTAATCAATAGCAAGATTTGG 300  
QY 358 AGATGAAAAATACAGATTTGTACACAAATGCTCGAATTTGGTGGAAAAATCGGCAAGACA 417  
DB 301 AGATGAAAAATTTAGAAATGTCACCCAGATGCTCGGATTTGGTGAGAAACCGATCGAGACA 360  
QY 418 AATGAGATTACACTACAGTGTTCACAGATCTCTGCTGAAAGTGAAACGAGCCTCAGATA 477  
DB 361 AATGAGCTGCATTACAGTGTTCACAGTGTTCAGAGATCTCTGCTGAAAGTGAGCGCCTCAGACA 420



Db 781 GGAGTCGTCATTCACAGTGTTCAGGATCCTGCTCGAAGAGTGAGCGGCTTCAGACAGTC 840  
Qy 481 AAGATGGATTCAGACCAACCAAGATCTTCAAGAGACCCCGCAGCGGACCGAG 540  
Db 841 GAAGATGGATTCAGATCAACCGGAAGATCTTCTAGAAGACCTCGAAGACAGAGGACCGAG 900  
Qy 541 TGAAGACCGTGATTTATGTGCATGCGCAATGGATTGAAGACTGTGATGATCAGCGCAC 600  
Db 901 TGAGACCGTGATTTATGTGCATGCGCAATGGATTGAAGACTGTGATGATCAGCGCAC 960  
Qy 601 TAAAGAAAAGAAATCAAGTCAGCAAGAAAAGAAACCGTCCAGGCCCAAGCAGGAAG 660  
Db 961 GAAAGAAAAGAGATCCAGTCCCGCAAGAGAAAGAGCGCTCCAGGCCCAAGCAGGAG 1020  
Qy 661 GGAAGCTTCACTGTTGAGTTTGCATAGATTCCTATGAACCTACATCTACTGCTTATGCAA 720  
Db 1021 GGAGGCATCCCTGTGAGTTTGCATAGATTCCTATGAACCTACATCTACTGCTTATGCAA 1080  
Qy 721 CCAAGTGCTTATGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
Db 1081 CCAAGTGCTTATGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Qy 781 TCATTTTTCATGTTTTCATTTACCTTACCTATTAACCAAGGGGAAATGTTATGCCCCAAAGTG 840  
Db 1141 TCATTTTTCATGTTTTCATTTACCTTACCTATTAACCAAGGGGAAATGTTATGCCCCAAAGTG 1200  
Qy 841 CAGGGAGATATGAGAAAACATGACAAAGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 1201 TAGGGAGACATGAGAAAACATGACAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
Qy 901 A 901  
Db 1261 A 1261

## RESULT 13

AC132966/c  
LOCUS AC132966 231311 bp DNA linear HTG 19-NOV-2002  
DEFINITION Rattus norvegicus clone CH230-210B24, WORKING DRAFT SEQUENCE, 2  
unordered pieces.  
AC132966  
VERSION AC132966.4 GI:25073360  
KEYWORDS HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 231311)

## REFERENCE

AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Huiy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, J., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensuhewa, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norrie, S., Nwaokelemeh, O., Okwionu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Popper, F., Poindecker, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajis, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, E., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE  
Unpublished

2 (bases 1 to 231311)

Rat Genome Sequencing Consortium.

## REFERENCE

Direct Submission

Submitted (05-SEP-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 231311)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23264819.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KB21

Center clone name: CH230-210B24

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 215506 bases at least Q40

Consensus quality: 218204 bases at least Q30

Consensus quality: 220107 bases at least Q20

Estimated insert size: 224360; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft.html).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 2 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 225634: contig of 225634 bp in length  
 \* 225635 225734: gap of unknown length  
 \* 225735 231311: contig of 5577 bp in length.

## FEATURES

## source

1. .231311  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-210B24"

3446. .102459

/notes="clone boundary"

clone\_end:T7

site:

end\_sequence: BH343917"

complement(218103. .218757)

/notes="clone boundary"

clone\_end:Sp6

site:

end\_sequence: BH343918"

220260. .222134

/notes="wgs end extension"

clone\_end:Sp6"

222185. .225634

/note="wgs end extension"

clone\_end:Sp6"

BASE COUNT 60763 a 49448 c 50231 g 60699 t 10170 others

## ORIGIN

Query Match 63.6%; Score 686.8; DB 2; Length 231311;  
 Best Local Similarity 90.8%; Pred. No. 1.3e-136;  
 Matches 744; Conservative 0; Mismatches 77; Indels 1; Gaps 1;

QY 233 TATCAAGAAACGTTAAAGGAAATGATGTTCTACGAAAATATAGAAGAGATGAT 292  
 DB 53296 TTTTGTAGAACTTTAAAGGAAATGATGTTCTATGAAAATATAGAAGAGATGAT 53237

QY 293 TTAACACAGAAAGAGCTTACAGAGCTTCTCCAGAGACATTAATTAATGACAGAA 352  
 DB 53236 TCAACACAGAAAGAGCTTACAGAGCTTCTCCAGAGACATTAATTAATGACAGAA 53177

QY 353 TTGGGAGATGAAAATACAGATGTTTACACAAATGCTCAAAATGTTGGGAAATCGGCA 412  
 DB 53176 TTGGGAGATGAAAATACAGATGTTTACACAAATGCTCAAAATGTTGGGAAATCGGCA 53117

QY 413 AGACAAATGGAGTTACACTCACAGTGTCTTCCAGATCTCTGCTGAAAAGTGAACGAGCTCA 472  
 DB 53116 AGACAAATGGAGTTGATTCATTCACAGTGTCTGCTGAAAAGTGAACGAGCTCA 53057

QY 473 GATAAGCAAGATGGATTCACGCCAACCAAGAGATCTTCAAGAGACCCCGAGCGAG 532  
 DB 53056 GACAAATCCAGATGATTCAGTCAACCGAAGATCTTCCAGAGACCTCGAAGACAG 52997

QY 533 CGGACAGTGAAGCCGTGATTTATGTCATCGGAAATGGGATGGAAGTGTGATGAT 592  
 DB 52996 CGGACAGTGAAGCCGTGATTTATGTCATCGGAAATGGGATGGAAGTGTGATGAT 52937

QY 593 CAGCCACCTAAAGAAAGAAATCCAGTCAGCAAGAAAGAAAGAAAGCTCCAGGCGCAAG 652  
 DB 52936 CAGCCACCTAAAGAAAGAAATCCAGTCAGCAAGAAAGAAAGCTCCAGGCGCAAG 52877

QY 653 CAGGAAAGGAGCTTCACTGTTGATTTGCAATAGATCTTAATGAACCTACACTGTC 712  
 DB 52876 CAGGAGAGGAGCGCTCACCTGTTGATTTGCAATAGATCTTAATGAACCTACACTGTC 52817

QY 713 TTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGGAAGTGAACAGTGTCCAAT 772  
 DB 52816 TTATGTAACCAAGTGTCTTATGGGAGATGATAGGCTGTGACAATGAACAGTGTCCAAT 52757

QY 773 GAATGTTTCACATTTTCTGTTTCTACCTATACCTATAAACCAAGGGGAATGTTATGC 832  
 DB 52756 GAGTGGTTTTCACATTTTCTGCGTTTCTACCTACCTATAAACCAAGGGGAATGTTATGC 52697

QY 833 CCNAAGTCAGGGGAGATATAGAAACCAATGACAAAGTACTGAAAACAAAGAAAG 892  
 DB 52696 CCNAAGTCAGGGGAGACAGTGAAGAAACCAATGACAAAGTACTGAAAACAAAGAAAG 52637

QY 893 GATAGAAAGTCCAGGCTAGTAAAGGCCATCCACATTTTAAAGGGTATTTCTCTTTTATAT 952  
 DB 52636 GAGAGAAAGCGAGGCTAGTGAAGGCCATCC-GGTTTAAAGGGTCTTTGCTCTTTATAT 52578

QY 953 AATTCGTTTCTTTCAGAAAATGTTTTCAGGTAATGCAATGCAATGCAATGCAATTTT 1012  
 DB 52577 AATTCGTTTCTTTCAGAAAATGTTTTCAGGTAATGCAATGCAATGCAATTTT 52518

QY 1013 AATCATTAGTATTAATGGTGTATTAAGTGTGTTGACTTTG 1054  
 DB 52517 AATCATTAGTATTAATGGTGTATTAAGTGTGTTGACTTTG 52476

## RESULT 14

## AC128348/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## 1 bases 1 to 245104)

## Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Amin, A., Angiano, D., Allen, C., Allen, H., Alsbrooks, S., Ayodeji, M., Baca, E., Baden, H., Anyalebechi, V., Aoyagi, A., Barber, M., Barnstead, M., Benahmed, F., Baldwin, K., Bandaranaike, D., Blyth, P., Brown, M., Biswal, O., Blair, J., Blankenburg, K., Burch, P., Burrell, K., Calderon, E., Bryant, N., Buhay, C., Carter, K., Cavazos, I., Ceasar, H., Chen, Z., Chu, J., Cardenas, V., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Chacko, J., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Dwyer, C., Ding, Y., Dinh, H., Divya, K., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, E., Geet, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Hawlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loshchukova, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Register, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,







Db 216959 TCAACACGAGAAAAACGCTCAGCAGCATCTCCAGAGAGCATTAATCAATAGCCAAAGAA 216900  
QY 353 TTGGGAGATGAAAAATACAGATTGTTACACAAATCTCGAATTTGGTGGAAATCGGCA 412  
Db 216899 TTGGGAGATGAAAAATCAGATTGTTACACAGATCTGGAATTTGGTGGAAAAACGAGCC 216840  
QY 413 AGACAAATGGAGTTACACTCACAGTGTCTTCCAAAGATCTGCTGAAAAGTGAACGAGCCTCA 472  
Db 216839 AGACAAATGGAGCTGCATTTCACAGTGTCTTCCAGGATCTGCTGAAAAGTGAACGAGCCTCA 216780  
QY 473 GATAAAGCAAGATGGATTCCAGCCAAACAGAAAGATCTTCAAGAAAGACCCCGCAGGCAG 532  
Db 216779 GACAAATCCAGATGATTCCAGTCAACCGGAAAGATCTTCCAGAAAGACCTCGAAGACAG 216720  
QY 533 CGGACCAAGTGAAGCCGCTGATTATGTGCATGGCAANTGGGATTTGAAGACTGTGATGAT 592  
Db 216719 CGGACCAAGTGAAGCCGCTGATTATGTGCATGACAAACGGGATTTGATGACTGCGATGAT 216660  
QY 593 CAGCCACCTTAAAGAAAGAAATCCAAAGTCAGCAAGAAAGAAACGCTCCAAAGGCCAAG 652  
Db 216659 CAGCCACCGAAGAAAGAAATCCAAATCCGCCAAGAAAGAAAGCGCTCCAAAGGCCAAG 216600  
QY 653 CAGGAAAGGGAAGCTTTCACCTGTGTGAGTTTGCATAGATCTTAATGAACCTACATACTGC 712  
Db 216599 CAGGAGAGGGAAGCGTCACTGTGTGAGTTTGCATAGATCTTCCAAACGAGCCCACTACTGC 216540  
QY 713 TTATGCAACCAAGTGTCTTATGGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAAT 772  
Db 216539 TTATGTAACCAAGTGTCTTACGGGGAGATGATAGGCTGTGACAAATGAACAGTGTCCAAT 216480  
QY 773 GAATGTTTTCACCTTTTCATGTGTTTTCACCTTACCTATAAACCAAGGGGAAATGGTATTGC 832  
Db 216479 GATGTTTTCACCTTTTCTCGCTTTCACCTACCTATAAACCAAGGGGAAATGGTATTGC 216420  
QY 833 CCAAGTGCAGGGGAGATTAATGAGAAAAACAATGGACAAAAGTACTGAAAAAGACAAAAAAG 892  
Db 216419 CCAAGTGCAGGGGAGACAGTGAAGAAACCATGACAAAAGTACTGAAAAGACAAAG 216360  
QY 893 GATAGAAGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTATTTGTCTTTTATAT 952  
Db 216359 GAGAGAAGAGCGAGGTAGTGAAGGCCATCC-GGTTTTAAAGGGTCTTTGTCTTTTATAT 216301  
QY 953 AATTGCTTTGCTTTTCAGAAAATGTTTTAGGGTAAATGCAAGACTATGCAATAATTTT 1012  
Db 216300 AATTGCTTTTACTTTTCAGAAAATGTTTTAGGGTAAATGCAAGACTATGCAATAATTTT 216241  
QY 1013 AATCATTAGTATTAAATGGTGTATTAAAGTTGTGTACTTTG 1054  
Db 216240 AATCATTAGTATTAAAGTTGTGTACTTTG 216199

Search completed: January 12, 2004, 18:30:25  
Job time : 3984 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004. Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2004, 19:31:18 ; Search time 281 Seconds  
(without alignments)

3460.255 Million cell updates/sec

Title: US-09-513-365A-1

Perfect score: 1481

Sequence: 1 MUGQQQQQLYSSAALLTGER.....DNEKTMDSKSTKDKRRSR 280

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2276164 seqs, 1736306516 residues

Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp  
-Q=/cgn2.1/USPTO.spool\_p/US09513365/runat\_12012004\_163934\_28802/app\_query.fasta.1.455  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2.6/prodata/2/pubna/US07\_PUBCOMB.seq:  
2: /cgn2.6/prodata/2/pubna/PCT\_NEW\_PUB.seq:  
3: /cgn2.6/prodata/2/pubna/US06\_NEW\_PUB.seq:  
4: /cgn2.6/prodata/2/pubna/US06\_PUBCOMB.seq:  
5: /cgn2.6/prodata/2/pubna/US07\_NEW\_PUB.seq:  
6: /cgn2.6/prodata/2/pubna/PCTUS\_PUBCOMB.seq:  
7: /cgn2.6/prodata/2/pubna/US08\_NEW\_PUB.seq:  
8: /cgn2.6/prodata/2/pubna/US08\_PUBCOMB.seq:  
9: /cgn2.6/prodata/2/pubna/US09A\_PUBCOMB.seq:  
10: /cgn2.6/prodata/2/pubna/US09B\_PUBCOMB.seq:  
11: /cgn2.6/prodata/2/pubna/US09C\_PUBCOMB.seq:  
12: /cgn2.6/prodata/2/pubna/US09\_NEW\_PUB.seq:  
13: /cgn2.6/prodata/2/pubna/US09\_NEW\_PUB.seq:  
14: /cgn2.6/prodata/2/pubna/US10A\_PUBCOMB.seq:  
15: /cgn2.6/prodata/2/pubna/US10B\_PUBCOMB.seq:  
16: /cgn2.6/prodata/2/pubna/US10\_NEW\_PUB.seq:  
17: /cgn2.6/prodata/2/pubna/US60\_NEW\_PUB.seq:  
18: /cgn2.6/prodata/2/pubna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-10-115-899-6  
; Sequence 6, Application US/10115899  
; Publication NO. US20020151025A1  
; GENERAL INFORMATION:  
; APPLICANT: Ostuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human ING1L gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/10/115,899  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/601,478  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens

1	1481	100.0	840	14	US-10-115-899-6	Sequence 6, Appli
2	1481	100.0	1078	14	US-10-115-899-7	Sequence 7, Appli
3	853.5	57.6	873	11	US-09-968-653A-4	Sequence 4, Appli
4	728	49.2	8487	10	US-09-764-877-3454	Sequence 3454, Ap
5	722	48.8	2061	11	US-09-968-653A-2	Sequence 2, Appli
6	662	44.7	633	11	US-09-968-653A-6	Sequence 6, Appli
7	474	32.0	421	11	US-09-918-995-37033	Sequence 37033, A
8	359	24.2	1864	13	US-10-117-722-435	Sequence 435, App
9	359	24.2	1864	15	US-10-037-270-435	Sequence 435, App
10	314	21.2	451	10	US-09-738-973-146	Sequence 146, App
11	314	21.2	451	10	US-09-854-133-146	Sequence 146, App
12	314	21.2	451	15	US-10-144-649A-146	Sequence 146, App
13	277.5	18.7	813	10	US-09-938-842A-1178	Sequence 1178, Ap
14	263.5	17.8	993	10	US-09-801-368-257	Sequence 257, App
15	253	17.1	897	13	US-10-032-585-6348	Sequence 6348, Ap
16	251.5	17.0	912	15	US-10-128-714-2502	Sequence 2502, Ap
17	223	15.1	1890	13	US-10-027-632-261924	Sequence 261924,
18	223	15.1	1890	13	US-10-027-632-261925	Sequence 261925,
19	223	15.1	1890	14	US-10-027-632-261924	Sequence 261924,
20	223	15.1	1890	14	US-10-027-632-261925	Sequence 261925,
21	219	14.8	1254	15	US-10-128-714-1502	Sequence 1502, Ap
22	219	14.8	1254	15	US-10-128-714-5502	Sequence 5502, Ap
23	219	14.8	1254	15	US-10-128-714-5502	Sequence 5502, Ap
24	219	14.8	3254	15	US-10-128-714-5502	Sequence 5502, App
25	219	14.8	3254	15	US-10-128-714-5502	Sequence 5502, App
26	186	12.6	36211	12	US-10-085-117-49	Sequence 49, Appl
27	160	10.8	456	11	US-09-918-995-16047	Sequence 16047, A
28	138	9.3	29346	12	US-10-085-117-52	Sequence 52, Appl
29	134	9.0	487	13	US-09-814-353-12868	Sequence 12868, A
30	134	9.0	2167	13	US-09-814-353-19129	Sequence 19129, A
31	133	9.0	508	13	US-09-814-353-88	Sequence 88, Appl
32	133	9.0	508	13	US-09-814-353-6483	Sequence 6483, Ap
33	130.5	8.8	7805	13	US-10-006-285-400	Sequence 400, App
34	130.5	8.8	7805	14	US-10-044-030-370	Sequence 370, App
35	129.5	8.7	915	10	US-09-764-864-471	Sequence 471, App
36	129.5	8.7	2320	15	US-10-175-523-2320	Sequence 2320, App
37	128.5	8.7	352	10	US-09-867-701-2399	Sequence 2399, Ap
38	128	8.6	65	13	US-09-908-975-23744	Sequence 23744, A
39	128	8.6	3391	13	US-10-172-094-96	Sequence 96, Appl
40	128	8.6	3391	15	US-10-072-094-96	Sequence 96, Appl
41	128	8.6	4467	13	US-10-172-094-88	Sequence 88, Appl
42	128	8.6	4467	15	US-10-072-094-88	Sequence 88, Appl
43	128	8.6	4895	13	US-10-115-831-134	Sequence 134, Appl
44	127.5	8.6	1820	9	US-09-737-178-17	Sequence 17, Appl
45	127.5	8.6	1820	9	US-09-737-178-37	Sequence 37, Appl

; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
US-10-115-899-6

```

Alignment Scores:
Pred. No.:      4.03e-159
Score:          1481.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    100.00%
DB:             14
Matches:        88
Conservative:   0
Mismatch:       0
Indels:         0
Gaps:           0
length:        88

```

US-09-513-365A-1 (1-280) x US-10-115-899-6 (1-840)

Qy	1	MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaIalaLeuLeuThrGlyGluArg	20
Db	1	ATGTAGGGCAGCAGCAGCAACTGTGTCGCGCGCGTCTCTGACCGGGAGCG	60
Qy	21	SerArgLeuLeuThrCystTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis	40
Db	61	AGCCGGCTGCTCACTGCTACTGTCGACGACCTACCTTGAGTGGCTGGAGTGCCTGCCCCAC	120
Qy	41	AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu	60
Db	121	GACATGCAGAGGAACTGTCTGTGTCGCGAGCTGGCAACAATAATCAAGAAACGTTTA	180
Qy	61	LysGluIleAspAspValTyrGluLysTyrIlyslsYsGluAspAspLeuAsnGlnLys	80
Db	181	AAGGAAATTTGATGTGTCTACGAAAAATATTAAGAAAGAATGATTTTAAACCAAGAGAA	240
Qy	81	ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys	100
Db	241	CGTCTACAGCAGCTTCTCCAGAGACCTAATTAATAGTCAAGNAATGGGAGATGAAAA	300
Qy	101	IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu	120
Db	301	ATACAGATTGTTACACAAATGCTCGAATTCGTGGAAAAATCGGCAGACAAATGGAGTTA	360
Qy	121	HisSerGlnCysPheGlnAspProHlaGluSerGluArgAlaSerAspLysAlaLysMet	140
Db	361	CACTCACTGTTTCAAGATCCTGCTGAAGTGAACGAGCTCAGATAAAGCAAAAGATG	420
Qy	141	AspSerSerGlnProGluArgSerSerArgArgProArgGlnArgThrSerGluSer	160
Db	421	GATTCCAGCCAACCGAAGATCTTCAAGAGAGACCCCGCAGGCAGCGACCACTGGAAGC	480
Qy	161	ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu	180
Db	481	CGTGATTATATGTCACATGCGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAA	540
Qy	181	LysLysSerLysSerAlaIlyslsYsLysArgSerLysAlaLysGlnGluArgGluAla	200
Db	541	AGAATACTCAAGTCAGCAAGAAAAAGAAACGCTCCAGGCCAAGCAGGAAAGGGAAGCT	600
Qy	201	SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal	220
Db	601	TCACCTGTTGAGTTTGCAATAGATCTTAATGAACCTACATACCTGCTTTATGCAACCAAGTG	660
Qy	221	SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluThrPheHisPhe	240
Db	661	TCTTATGGGGAGATGATAGGATGTGCACAAATGAACAGATGTCCAAATTTGAATGGTTTCACTTT	720
Qy	241	SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly	260
Db	721	TCATGTGTTTCACTTACCTATTAACCAAGGGGAATGGTATTGCCCAAGTGCAGGGGA	780
Qy	261	AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg	280
Db	781	GATAATGAGAAACATGCAAAAGTACTGAAAGACAAATAAGGATAGAGATCGAGG	840

## RESULT 2

US-10-115-899-7

US-10-113-899-7  
; Sequence 7, Application US/10115899









```

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37033
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(421)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-37033

```

Alignment Scores:		
Pred. No.:	9,48e-45	421
Score:	474.00	Matches: 85
Percent Similarity:	95.00%	Conservative: 5
Best Local Similarity:	80.00%	Mismatches: 10
Query Match:	32.01%	Indels: 0
DB:	11	Gaps: 0

US-09-513-365A-1 (1-280) x US-09-918-995-37033 (1-421)

181	LysLysSerLysSerAlaLysLysLysLysLysArgSerLysAlaLysLysGlnGluArgGluAla	200
Qy		
3	AAGAATTCTCTAGTCAGCAAAAGAAATAGAAACTCTCCGAGGCCACGCAGGAAAGGGAGCT	62
Db		
201	SerProValGluPheAlaIleAspProAsnGluProThrTyrcysLysCysAsnGlnVal	220
Qy		
63	TCACCTGTTGAGTTTGCAATAGATCCTTAATGAACCTACACTACTGCTTATGCAACCAAGTG	122
Db		
221	SerTyrglyGluMetileglyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe	240
Qy		
123	TCTTATGGGGAGATGATAGGATGTGACAATGAACACCGTCCAAATTGATGAGTGATTCATTT	182
Db		
241	SerCysValSerLeuThrTyrlsLysProLysGlyLysTrpTyrcysProLysCysArgGly	260
Qy		
183	TTATGTGTTTCACTTACCTATAAACCAAGGGGAATGGTATGCCCACACAGTGCAGGGA	242
Db		
261	AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg	280
Qy		
243	GATATATGAGAAAACAAATGGACCAAGTACTCATTAAGACAAAACACGATAGAAGATCGAGG	302
Db		

## RESULT 8

```

US-10-117-722-435
; Sequence 435, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 435
; LENGTH: 1864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(1420)
US-10-117-722-435

```

Alignment Scores:			
Pred. No.:	1.03e-30	Length:	1864
Score:	359.00	Matches:	102
Percent Similarity:	34.22%	Conservative:	53
Best Local Similarity:	22.52%	Mismatches:	100
Query Match:	24.24%	Indels:	198
DB:	13	Gaps:	11

  

US-09-513-365A-1 (1-280) x US-10-117-722-435 (1-1864)			
Qy	22	ArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHisasp	41
Db	155	AGGCCCGGATGTTGTACCTAGAGACTATCTGGAAATGATTGACAGCTTCCTATGGAT	214
Qy	42	MetGlnArgAsnValSerValLeuArgLulLeuAspAsnLysTyrGlnGluThrLeuLys	61
Db	215	CTGGGAGCGCTTCACGGAAATGCGGAGATGACCTGCAGGTGCAGGTGCAGAAATGCAATG	271
Qy	62	GluIleAspAspValTyrGluLysTyrLysGluAspAspLeuAsnGlnLysLys	80
Db	272	-----GATCAACTAGACAAGAAGTCAGTGAATTCCTTTATGAATGCCAAGAAAAAT	322
Qy	81	-----ArgLeuGlnGlnLeu-----LeuGlnArgAlaLeuIleAsnSerGln	94
Db	323	AAACCTGAGTGGAGGAGCAATGGCATCCATCAAAAAGACTACTATAAAGCTTTG	382
Qy	95	GluLeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArg	114
Db	383	GAAGATGCAGATGAGAAGTTTCAGTTGGCAACACAGATATATGACTTGGTAGATCGACAC	442
Qy	115	AlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAla	134
Db	443	TTGAGAAGCTGGATCAGGAACCTGGCTAAAGTTTAAATGGAGCTGGAAGCTGATTAATGCT	502
Qy	134	-----	134
Db	503	GGAATTACAGNAATATTAGAGAGCGGATCTTTGGNAATTAGACACTCTCTTCACAGCCAGTG	562
Qy	134	-----	134
Db	563	AACAATCACCATGCTCATTCACATACTCCAGTGGAAAAAAGGAATATAATCAACTTCT	622
Qy	134	-----	134
Db	623	CACCATACGACAAACAGATCATATTCTCGAAAAAGAAATTTAAATCTGAAGCTCTTCTATCC	682
Qy	135	-----SerAspLysAlaLys-----	139
Db	683	ACCTTACGTCAGATGCTCTAAGGAAATAACATAGTGTTCGAAATAATAATTCACCA	742
Qy	140	-----MetAspSerSerGlnPro-----	145
Db	743	GCCTCTTTAACAATGCCTACAATGTGAATTCCTCCCAACCTCTGGGATCTCTATAACATT	802
Qy	145	-----	145
Db	803	GGCTGTTATCTTCAGGAACCTGGTGAGGGGCAATTACATGGCAGCTGCTCAAGCAGTT	862
Qy	146	GluArgSerSerArgProArgArgGlnArgThrSerGlu-----	159
Db	863	CAGCTACAGCTCAGATCAAGGAGGAGCAAGACATCAAGTTTAAAGCCAGTTATGAA	922
Qy	159	-----	159
Db	923	GCATTTAAGATAATAGACTTTTCAGTTGGGAAAAAGAAATTTTCAATGGCCAGGAAAAACAGTT	982
Qy	160	-----SerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnPro	177
Db	983	GGCTATTATCATCTCTGGGCACTTATGACAACATTAAACAGAAATGCCAGTTTCATCAGCA	104
Qy	178	ProLysGluLysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGlu	197





```

Db      256  ACGGCGGCGAGGTGGACAGCCACGTGTGAGAGCTGTTTCGAGGGCGCAGCAGAGCTGGGCGACACA 315
      |||||:::||||:::||||:::||||:::||||:::
Qy      134  AlaSerAspLysAlaLysMetAspSerGlnPro----- 145
      |||:::||||:::
Db      316  GCGGGCAACAGCGGCAAGGCTGGCGCGACAGGCCCAAAAGGCGAGCGSCAGCGCAGGCT 375
      |||:::||||:::
Qy      146  ---GluArgSerArgArgProArgGlnArgGlnThrSerGluSerArgAsp 162
      :::::|||||:::|||||:::|||||:::
Db      376  GACAAGCCCAACAGCAAGCGCTCACGGCGGCGAGCGCAACACAGAAACCGTGAG 429
      :::::|||||:::|||||:::|||||:::
RESULT 12
US-10-144-649A-146
; Sequence 146, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 146
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-144-649A-146

```

US-09-938-842A-1178  
; Sequence 1178, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPL300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1178  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1178

Alignment Scores:  
Pred. No.: 6,13e-22 Length: 813  
Score: 277.50 Matches: 77  
Percent Similarity: 45.35% Conservative: 40  
Best Local Similarity: 29.84% Mismatches: 81  
Query Match: 18.74% Indels: 61  
DB: 10 Gaps: 9

US-09-513-365A-1 (1-280) x US-09-938-842A-1178 (1-813)

Qy 27 TyrValGlnAspTyrLeuGluCysValGluSerLeuProHisAspMetGlnArgAsnVal 46  
Db 25 TATGTTGATGATTACCTTGAGTATGCAAGCACCTTCCCTGCAGAGCTACAAGATTACTA 84  
Qy 47 SerValLeuArgGluLeuAsp----- 53  
Db 85 AACACAGTTCGCGAACTAGACGAGAGATCTCAATGCTTGTTCACAGAGTTCATAAAG 144  
Qy 54 -----AsnLysTyrGlnGluThrLeuLysGluIleAspAspValTyrGluLysTyr 70  
Db 145 GGTAAATGGTAAATCATTAACAATAATGTTGTTGATGAGAGAGAAACGATTGAGAAATG 204  
Qy 71 LysLysGluAspAspLeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeu 90  
Db 205 CGTAAAGAGATTGAGTCTAGTACGAA----- 231  
Qy 91 IleAsnSerGlnGluLeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeu 110  
Db 232 ---AATGCGTTAAGTTGTGTACCGAAGAGTTTATTGCGCCGACAGCGTATGATCTT 288  
Qy 111 ValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu 130  
Db 289 ATAGATAGTCATGTAACACGACTTGATGAAGATCTCGAATAATTTT-----GCAGAA 339  
Qy 131 SerGluArgAlaSerAspLysAlaLysMetAspSerSerGln----- 144  
Db 340 GATTTAAAGCAAGAGGAAAAATTCACACGACGAGCCCTCTGTCTTCTCCTCCACTACCT 399  
Qy 145 -----ProGluArgSerSerArgArg-ProArgArgGlnArgThrSerGluSerArgAs 162  
Db 400 ATAGATTCTTAACCGGAAAGCGTAAGTCTCTATGCGCACACCTCAGCCCAAGAAGATT 459  
Qy 162 pLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGln-----ProPolysG 180  
Db 460 GATTATAGATAGACATGGGATCGT-GACAGGATTTTGACTCATGCTCTCTCTCTCTCT 513  
Qy 180 uLysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAl 200

Db 514 -----CCAGGAGCAATCGGAAGACCT 536  
 Qy 200 aserProVal---GluPheAlaIleaspProAsnGluProThrTyrCysLeuCysAsnG1 219  
 Db 537 TATGCAATTGAAGAGCAGCCCAATCGATCAAAACGAAACCACTTACTGTGTGCGCATCA 596  
 Qy 219 nValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysPro---IleGluTrpPh 238  
 Db 597 GTGTGCTTTGAGACATGATTCCTGTGACATGAGAAATGCCAAGGAGGTGAATGGTT 656  
 Qy 238 ehispheSerCysValSerLeuThrTyrLysPro-----LysGlyLysTrp 253  
 Db 657 TCACATATACATCGGTGGCTTCACACCTGAGACCAGATTCAAAGGGAAATGG 708

## RESULT 14

US-09-801-368-257

; Sequence 257, Application US/09801368

; Patent No. US20020128250A1

; GENERAL INFORMATION:

; APPLICANT: Bueby, Robert

; APPLICANT: Calli, Brian

; APPLICANT: Hecht, Peter

; APPLICANT: Holtzman, Doug

; APPLICANT: Madden, Kevin

; APPLICANT: Maxon, Mary

; APPLICANT: Milne, Todd

; APPLICANT: No. US20020128250A1man, Thea

; APPLICANT: Royer, John

; APPLICANT: Salama, Sofie

; APPLICANT: Sherman, Amir

; APPLICANT: Silva, Jeff

; APPLICANT: Summers, Eric

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

; FILE REFERENCE: 109272.147

; CURRENT APPLICATION NUMBER: US/09/801,368

; PRIOR FILING DATE: 2001-03-07

; PRIOR APPLICATION NUMBER: US 09/487,558

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: US 60/160,587

; PRIOR FILING DATE: 1999-10-20

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 257

; LENGTH: 993

; TYPE: DNA

; ORGANISM: Saccharomyces cerevisiae

US-09-801-368-257

Alignment Scores:

Pred. No.:	3.22e-20	Length:	993
Score:	263.50	Matches:	75
Percent Similarity:	37.04%	Conservative:	45
Best Local Similarity:	23.15%	Mismatches:	105
Query Match:	17.79%	Indels:	99
DB:	10	Gaps:	8

US-09-513-365A-1 (1-280) x US-09-801-368-257 (1-993)

Qy 28 ValGlnAspTyrLeuGluCysValGluSerLeuProHisAspMetGlnArgAsnValSer 47  
 Db 31 TTAACGACATTAAGTGTGGAGGAAATCCCACTGGCCACCTCCAGATATTAAAC 90  
 Qy 48 ValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeuLysGluLeuAspAspValTyr 67  
 Db 91 TTAACACGAAATAGATGCAAAATGTGTGATTCATGCGGAATTTGAACGAGAGGATA 150  
 Qy 68 GluLysTyrLysLysGluAspAsp-----LeuAsn 77  
 Db 151 GATAAGTTCTTGAAGAAAGACTTCAATAAAGATCAACAAACAAAGTAAGACTGCTCAAT 210  
 Qy 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGly 97

Db 211 AATATCAACAAGATTATTATGAAGAACTGATGCCATCGCTG----- 249  
 Qy 98 AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln 117  
 Db 250 GAGGAGAAATGTCATGTCTCATCTTATGCTGGATAATCTAGACAGATTGAGTCCCGG 309  
 Qy 118 MetGluLeu----- 120  
 Db 310 TTAGAAATGGCGTATGAAGTCGCAATCAAGAACACAGAAATCTTCTAGAGGTTTAAGACTG 369  
 Qy 121 -----HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAsp 136  
 Db 370 GGTGTGGACAACCATCCAGCAATCCACCTCCATCATGAACCTAATGAAAAAATAGAGAGC 429  
 Qy 137 LysAlaLysMetAspSerSerGlnProGluArgSer---SerArgArgPro----- 152  
 Db 430 AAATCAAAACAGCAAAATCGTCGAGGCACTGAAGAGCGAATCAAGAAAGAGGAGGATGGCT 489  
 Qy 153 -----ArgArgGln-----ArgThrSerGluSerArg 161  
 Db 490 GCCAACAGGAGGCGAGCGCAACATTACTCCGCCACGACACACCAACAGACGACTCAAG 549  
 Qy 162 AspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLys 181  
 Db 550 AACGACGCAAACTACCGAGGCGAGGAGGAGCAAGCAACACTGTTAACAACACACA 609  
 Qy 182 LysSerLys-----SerAlaLysLys 188  
 Db 610 AACTCAAGAAAAAGAGCCAAAGCTGCCAATACCAACAGCGCGATCCAGAGACCAAAAA 669  
 Qy 189 LysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIle--- 207  
 Db 670 CGCAAGAGGAGAGTTGCCACCACACGCCGTTTCCACCAAGCACTATCAGCAGCGCAACTGCC 729  
 Qy 207 ----- 207  
 Db 730 GTCAATAATGGCAGGATAGGTATACATCTACAGCGCTCAGGGGAGTTAGACGCTCGGAAAC 789  
 Qy 208 -----AspProAsnGluProThrTyrCys 215  
 Db 790 AGCAACAACAGCAGGATATCAAGACCAAAACCAACGACTACGGGGAACCGCTCTACTGC 849  
 Qy 216 LeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIle 235  
 Db 850 TACTGTACCAAGTGGCATACGGGGAATGGTGGGGTGTGATGGCGCAGACTGTGAGCTA 909  
 Qy 236 GluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTrpTyrCys 255  
 Db 910 GAATGGTTCCATTGCGCATGTATTGGACTCGAAACTCTACCTAAGGGCAAGTGGTATTGC 969  
 Qy 256 ProLysCysArg 259  
 Db 970 GACGACTGCAAA 981  
 RESULT 15  
 US-10-032-585-6348  
 ; Sequence 6348, Application US/10032585  
 ; Publication No. US20030180953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Terry, Roemer D.  
 ; APPLICANT: Bo, Jiang  
 ; APPLICANT: Charles, Boone  
 ; APPLICANT: Howard, Buseby  
 ; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
 ; FILE REFERENCE: 10182-005-999  
 ; CURRENT APPLICATION NUMBER: US/10/032,585  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 8000  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 6348  
 ; LENGTH: 897  
 ; TYPE: DNA  
 ; ORGANISM: Candida albicans

US-10-032-585-6348

Alignment Scores:

Pred. No.: 4,38e-19 Length: 897  
Score: 253.00 Matches: 84  
Percent Similarity: 36.25% Conservative: 36  
Best Local Similarity: 25.38% Mismatches: 108  
Query Match: 17.08% Indels: 104  
DB: 13 Gaps: 10

US-09-513-365A-1 (1-280) x US-10-032-585-6348 (1-897)

QY 19 GluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeu 38  
Db 4 GATACATCACTGTACTTGAGAAATACACACAGGAT-----TTATCTAATCTA 51  
QY 39 ProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGlu 58  
Db 52 CCACCTTGAGTGAGGCAATTTATTAGAGAAATCAAAAGCAAGATGTACAAAGTCTCAGAA 111  
QY 59 ThrLeuLys-----GluLeAspAsp 65  
Db 112 GCAGGAACGATACCAAACTCGTGACCATCACTTCATAAATTTATACGAACAAATGGG 171  
QY 66 ValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLysArgLeuGlnLeu 85  
Db 172 ACACCTAACTAAACATCCAAAGGAAGACCACTGTGATTCGAAG-----213  
QY 86 LeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGluAspGluLysLysLeuGlnLeuValThr 105  
Db 214 ATCGAAGAGATATGAAGCTTGTGCAGAAACTACAGAAAGAGAAATTTCTATTGGCTAAC 273  
QY 106 GlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPhe 125  
Db 274 ACTGCATTTGTTTGATATCGAAACATCTATATATTTTTCAG-----315  
QY 126 GlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnPro 145  
Db 316 -----ACAGACATTCGCAAGCTAGAAAGAGATGAATT- 347  
QY 146 GluArgSerSerArgArgProArgGlnArgThrSer-GluSerArgAspLeuCysHi 165  
Db 348 GTTGCTCTCTTTAGAACACCCCTCGAGCTAACTGAAGTATCTAAAGACGAATACGCAAA 407  
QY 165 sMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSerLys 185  
Db 408 ATCACTTAATGGCTTTTCTGATAGTCTCTGCCACACCAACCGCTAGGAATGGATCATC 467  
QY 185 rAla-----LysLysLys-----189  
Db 468 GGCACTCCAGTGCAGAAACAGTGAAGAAAGATCCAAAGAGAAACTTAGTGTAAAGG 527  
QY 190 -----LysArgSerLysAl 194  
Db 528 GGGGTCTCTTCTCTGCTCAGTCATCTCTGCTTCAAGACAGGTGAAGCGACTCCGGTC 587  
QY 194 aLysGlnGluArgGluAlaSerProValGlu-----204  
Db 588 AGAAGAGATAGAGGATCCTCTACCGTATGAGGAGGGGTCAATGGCGTTCAATGGAATGT 647  
QY 205 -----PheAlaIleAspProAsnGluProThr-----213  
Db 648 AGCAATGAGCATAAATAGTCAGCAGATGCAAAATGTCCAAACGGGGAAGATGCTGATAA 707  
QY 214 -----TyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnG 231  
Db 708 TAATCTATATTTGTTTGTCAACGTGTTTCAATTTGGTGAATCATTTGTTGTTGTAATGA 767  
QY 231 uGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysG 251  
Db 768 AGACTGTAACTATGAATGGTTTTCATTGGAGTTGTTGGGATCACAATCACTCTCTAAAGA 827  
QY 251 YLys-----TrpTyrCysProLysCysArgGlyAspAsnGluLysThrMetAspLys 269

Db 828 TGACGAAATTTGGTATTGTCTCTGATTGTCT-----TCAAA 863  
QY 269 rThrGluLysThrLysLysAspArgSer 279  
Db 864 GATGGAGAAAAGAAAAAGAAACGGGAAGAAC 894

Search completed: January 12, 2004, 19:54:49  
Job time : 290 secs

